(2)

RECOMMENDATIONS FOR IMPROVING AIR FORCE NONDESTRUCTIVE INSPECTION TECHNICIAN PROFICIENCY

FINAL REPORT

SwRI Project No. 17-7958-845
Contract No. DLA900-84-C-0910

Prepared for:

Air Force NDI Program Office
San Antonio Air Logistics Center
Air Force Logistics Command
Kelly AFB, Texas 78241-5000

December 20, 1988

SOUTHWEST RESEARCH INSTITUTE
SAN ANTONIO HOUSTON

SOUTHWEST RESEARCH INSTITUTE
Post Office Drawer 28510, 6220 Culebra Road
San Antonio, Texas 78284

RECOMMENDATIONS FOR IMPROVING AIR FORCE NONDESTRUCTIVE INSPECTION TECHNICIAN PROFICIENCY

FINAL REPORT

SwRI Project No. 17-7958-845
Contract No. DLA900-84-C-0910

Prepared by:

James E. Schroeder
Donald W. Dunavant
Joel G. Godwin

Prepared for:

Air Force NDI Program Office
San Antonio Air Logistics Center
Air Force Logistics Command
Kelly AFB, Texas 78241-5000

December 20, 1988

Approved:

Dennis W. Gilstad, Director
Department of Bioengineering

89 4 03 011

SUMMARY

The purpose of this research was to provide recommendations for improving the proficiency of Air Force nondestructive inspection (NDI) technicians. The work was conducted in three phases. In Phase I, areas of concern were identified from the existing literature, from discussions with NDI personnel, and from observations resulting from site visits (Laughlin and Chanute Air Force Bases). Each of the 277 identified concerns was classified into one of 11 categories of variables known to affect human performance. They were consolidated and condensed to form 51 overriding concerns. Resulting concerns were rated for criticality. In Phase II, 227 possible solutions were identified or created. In Phase III, possible solutions were used as elements of recommendations which were then evaluated by an expert panel. A modified Delphi procedure was utilized to determine estimates of promise, feasibility, and cost. Resulting ratings were analyzed to provide measures of overall and relative merit.

TABLE OF CONTENTS

I. BACKGROUND

In recent years, considerable concern has been expressed over the proficiency level of nondestructive inspection (NDI) technicians in the Air Force, other government agencies, and industry. The general concern is that the human technician is not performing at the level commonly assumed nor at a level commensurate with the capabilities of the equipment used. The Air Force's concern originated partly from the results of an extensive evaluation of NDI technician proficiency in the mid 1970's, commonly known as "Have Cracks - Will Travel." The main finding of that research was that the level of Air Force NDI proficiency was below existing requirements (see Lewis, Sproat, Dodd, and Hamilton, 1978).

Although Lewis et al. reported a great deal of variability among the technicians, the mean probabilities of detecting fatigue cracks were "at least 25 percent less than the previously assumed values." In addition to the low proficiency results, Lewis et al. reported other distressing findings. Specifically, variables expected to be related to NDI technician proficiency (i.e., formal education, age, classification, skill level, NDI experience and NDI training) accounted for very little of the variance in NDI performance.

Subsequently, a number of investigators have voiced concerns over Air Force NDI proficiency (specific Air-Force related references are identified in the section that describes areas of concern). It is important to emphasize that the concern for poor performance in NDI tasks has not been limited to the Air Force technician population. There has been a constant stream of articles addressing poor nondestructive evaluation (NDE) performance in the power industry (specific industry-related references are identified in the section that describes areas of concern).

In September of 1987, Southwest Research Institute (SwRI) was contracted to investigate what might be done to increase NDI technician proficiency. The major goal of this effort was to make recommendations for a program for increasing Air Force NDI technician proficiency. Because the tasks, skills, and procedures for Air Force NDI technicians are generally the same as for NDI technicians in other sectors, the general literature was also investigated. Subsequently, issues which were judged irrelevant for the Air Force NDI technician population (e.g., problems involving intergranular stress corrosion cracking in power plants) were eliminated from further consideration.

The approach taken was to (a) identify relevant areas of concern that could negatively impact the proficiency of Air Force NDI technicians; (b) seek possible solutions for identified

concerns; and (c) blend the most feasible, promising, and cost-effective potential solutions into recommendations for improving Air Force NDI technician proficiency. The following sections provide more details about the methodology used; identified areas of concern; identified possible solutions; and resulting recommendations.

II. METHOD

The Method section is divided into three subsections, corresponding to the three major phases of the effort: (a) identify areas of concern with regard to Air Force NDI technician performance; (b) document or create possible solutions for identified areas of concern; and (c) extract recommendations along with corresponding estimates of their promise, feasibility and cost.

A. Phase I: Identify Areas of Concern

The purpose of Phase I was to identify possible areas of concern that could negatively impact NDI technician proficiency. Areas of concern came from four sources: the existing literature dealing with Air Force NDI technician proficiency; the existing literature dealing with industrial NDE technician proficiency; comments generated by Air Force personnel during interviews and other coordination activities; and comments generated by the SwRI investigators from observations, interviews, site visits, and other coordination activities. Site visits included trips to the NDI laboratory at Laughlin Air Force Base and the NDI training facility at Chanute Air Force Base.

Initially, all identified areas of concern were documented, along with their associated source. A total of 277 areas of concern were initially identified. Because the main purpose of this exercise was to document areas of concern from different sources, there was substantial redundancy among the concerns. No effort was made to identify or document the first person, team, or organization to identify a specific concern. Rather, concerns were documented in the order they were identified. The initial list of concerns was then classified according to their strongest relevance to 11 general human-performance categories:

* Selection
* Initial Training
* On-the-Job Training
* Work Environment
* Policy
* General Performance Variables
* Procedures

* Motivation
* System
* Measurement
* Equipment

Next, areas of concern judged to be irrelevant for Air Force NDI were removed. Eliminated concerns generally dealt with either intergranular stress corrosion cracking or power industry certification issues.

Because there was overlap among the various sources, the remaining 230 concerns were condensed to 51 relatively specific concerns and three overriding issues.

Finally, to provide the sponsor with an indication of the criticality of each concern, the three primary SwRI investigators independently rated the 54 areas of concern for perceived criticality. For each area of concern, the project team members rated both criticality and their own expertise in rating that particular item. The resulting mean ratings of criticality were weighted according to each rater's evaluation of personal expertise.

The final list of concerns was presented to the Air Force NDI Program Office for review. They then provided feedback about which areas had higher priority.

B. Phase II: Identify Possible Solutions

Potential solutions for identified concerns were collected from the same four sources: the existing literature dealing with Air Force NDI technician proficiency; the existing literature dealing with industrial NDI technician proficiency; candidate solutions generated by Air Force personnel during interviews and other coordination activities; and candidate solutions generated by the SwRI investigators from observations, interviews, and other coordination activities. For each possible solution, the associated reference source was documented. Again, there was no attempt to document who was responsible for first offering the solution. All 227 candidate solutions were cross-referenced to the relevant area(s) of concern. Possible solutions were also reviewed by the NDI Program Office.

C. Phase III: Generate and Evaluate Recommendations for Improving Air Force NDI Technician Proficiency

In the last phase, possible solutions were combined and condensed to form one or more recommendations for each area of concern. These were candidates for the final recommendations.

To provide an estimate of the efficacy of the derived recommendations, a panel of experts was used. The purpose of this exercise was to achieve a group estimate of the potential promise, feasibility, and cost of each recommendation. In addition, the group worked to clarify the meaning of any ambiguous recommendations.

The panel comprised three project team members from SwRI and three senior subject-matter experts from the Air Force NDI Program Office. A modified Delphi approach was taken. An expert in decision support, not familiar with the project, was brought in to act as an impartial moderator. A total of 75 candidate recommendations were addressed (however, some redundancy occurred since some of the recommendations were evaluated relative to more than one area of concern). The moderator presented a brief discussion of some common errors that experts often commit in such exercises (e.g., snap judgement, succumbing to recency effects, etc.).

On the first pass, the moderator read the relevant area of concern as well as the recommendation. There was no group discussion at that time, although panelists were urged to seek clarification about any ambiguous wording. Next, panelists were allowed two minutes of silence to think about the recommendation. They were asked to rate the recommendation on potential promise using a scale of one (not promising) to five (very promising). They were then asked to rate their confidence in their rating of potential promise on a scale of one (not confident) to five (very confident). Panelists were next asked to rate the recommendation for feasibility using a scale of one (not feasible) to five (very feasible). Similarly, they were asked to rate their confidence in the feasibility rating. Next, they were asked to rate the financial cost of following the recommendation on a scale of one (very high) to five (very low). Finally, they were asked to write down any changes they would make to the stated recommendation to clarify or improve the intent.

Following the first pass, resulting rating data were analyzed. For each recommendation, a mean rating was determined for potential promise, feasibility, and cost. In addition, recommendations with substantial disagreement were noted. Due to the speculative nature of the rating task, a liberal criterion for disagreement was adopted. Specifically, if the ratings for promise, feasibility or cost exceeded a range of three on the 5-point scale, then that item was flagged to be reassessed in the second pass because of disagreement. Ranges less than three were judged to constitute adequate agreement.

The panel was then reconvened to discuss the issues and attempt to increase agreement about those recommendations which resulted in varied reactions during the first pass. In that panel session, the disagreements were addressed in the following

general order: disagreements about promise followed by disagreements about feasibility. Disagreements about cost were not reassessed because (a) assessment of cost was expected to vary more because the cost estimates depend on the method of implementation and (b) time limitations because of conflicting schedules of the six experts. Also, recommendations were addressed in the opposite of the initial temporal order to counterbalance any systematic cumulative effects.

In the second pass, the panelists first read the disputed recommendation. Panelists were provided the distribution of ratings from the first pass. Open discussion was then encouraged to resolve any ambiguities or differences of opinion. At the end of discussion or 4 min. (whichever came first), discussion was closed and all panelists again rated the item.

Final data analyses were conducted on the ratings obtained. Disagreements (ratings with a range greater than three) were noted, but no further attempts to increase agreement were made. Finally, recommendations were rank ordered within each general category according to the sum of their ratings for promise and feasibility. Aggregate ratings for promise, feasibility, and cost were also computed and reported.

III. RESULTS OF PHASE I: AREAS OF CONCERN

A. Identified Areas of Concern and Sources

In this section, initial areas of concern are listed. These were found in the technical literature or through observations. They have been classified according to 11 human performance categories. Within each category, concerns were further divided into those which specifically addressed Air Force NDI issues (AF NDI) and those which specifically addressed industrial nondestructive evaluation issues (GENERAL NDE). In the following list, each concern is identified by an alphanumeric code. Following the code, an abbreviated reference is noted. More detailed reference information can be found in the References section. After each concern, a reference number is provided to allow the reader to cross-reference the original area of concern (found in this section) to the appropriate condensed area of concern (found in the next section). At the end of this list, a final section presents those concerns which were identified, but were judged to be irrelevant for Air Force NDI.

SELECTION:

AIR FORCE NDI

A-1. SwRI observation: In a field involving visuospatial information, there is little attention paid to technician selection based on spatial ability. 2.1

A-2. Summers: Reported no correlation between formal education and performance (however, he later reported those with more than 2 yrs college performed better). 2.2

A-3. Lewis et al. (1978): There is minimal influence due to formal education , technician's age, classification, skill level, NDI experience and NDI training. 2.2

A-4. SwRI - Chanute visit, Instructor's comment: NDI should be in electronics career field, not general career field. 2.1

A-5. SwRI - ALC/ATC interactions: NDI should be in electronics general career field instead of general. 2.1

A-6. Summers: Volunteers for NDI were no better than non-volunteers. 2.2

A-7. Summers: Persons with airframe or metal experience did no better than those without. 2.2

SELECTION:

GENERAL NDE

B-1. Karimi: We are not selecting the best people. The people we are getting cannot endure "physical hardships" of plants. 2.1

B-2. Wheeler et al.: The following were considered to be important attributes but are not used to select: ability to concentrate (11 of 12); understanding of NDT theory (10 of 12); patience (9 of 12); tolerance of environmental conditions (9 of 12); manual dexterity (7 of 12); and mathematical ability (7 of 12). 2.2

B-3. Davies and Tune: They found no difference due to intelligence or gender in NDT performance. 2.2

B-4. Spanner et al.: "No research results on the effects of individual difference variables on NDT performance could be found in the literature." 2.2

B-5. Triggs et al.: No research has been carried out in individual differences in NDE performance. 2.2

B-6. Michalski et al.: Much NDE is visually oriented - yet many people do not know of their eye weaknesses, many have wrong lenses, and some don't take glasses to work. There are undetected eye diseases. Medications and alcohol influence the eyes and can also increase fatigue. Authors list 9 example eye problems that could affect NDE proficiency. 2.3

B-7. Yonemura: There are known within- and between-subject differences and between-facility differences in visual performance. 2.3

B-8. Wheeler et al.: Reported that 4 of 12 felt that ability to work with others was critical trait of good NDI technician. 2.2

RESIDENT OR INITIAL TRAINING:

AIR FORCE NDI

C-1. SwRI observation: The NDI personnel do not get enough theory to become technicians, possibly more theory than necessary to become operators. 1.2, 3.1

C-2. SwRI observation: There is a danger in selecting trainers for class "n+1" from class "n," especially when no additional training or field experience occurs between the time the individual is a student and a trainer. Over generations, misinformation can be introduced and be perpetuated without detection. 4.1

C-3. Summers: There was no significant correlation between amount of NDI training and eddy or ultrasonic performance. 4.9

C-4. Summers: The ratings of OJT and resident training were not correlated with performance. (SwRI observation: yet, Chanute uses ratings as primary indicator that something is wrong.) 4.9

C-5. SwRI - from Lewis et al.: There were differences in performance far beyond those of technician differences; i.e., there were some cracks that were just difficult to detect, regardless of "length." This is a problem for their assumption that length is <u>the</u> relevant difficulty variable: apparently no one has conducted an error analysis to discover (other than length) what factors determine difficulty, ways of using the equipment to catch such flaws (e.g., perhaps identifying certain subtle signal patterns), and then specifically fed back such information to the training community so that they could <u>train</u> on those methods? 4.8

C-6. SwRI - from Lewis et al.: Hours of training seemed to affect performance in ultrasound but not other techniques. 4.9

C-7. Lewis et al. (1978): An inverse relationship between training and performance was reported! 4.9

C-8. SwRI - from Lewis et al.: They reported that logistics centers training is no problem because civilians are involved and there is low turnover; however, there is no evidence that this was the case. 4.9

C-9. Eagle, 1985: No computer-based training is now available for the initial training course. 4.4, 4.5

C-10. Eagle, 1985: They concluded that "... comprehension and retention of concepts and theory are the most difficult part of NDI training." 3.1, 4.2

C-11. Eagle, 1985: The computer-based trainer proposed by Human Resources Laboratory (HRL) only covers one rarely utilized procedure and does not address "other critical concepts." 4.8

C-12. Eagle, 1985: There is a need to "increase understanding of complex, invisible inspection principles and concepts." 4.5

C-13. Eagle, 1985: Currently, there are -3 level students who could learn more than "the minimum." 4.4

C-14. SwRI - Laughlin visit: Some technicians indicated that the initial training at Chanute was good but could use more hands-on components and more emphasis on some of the techniques (e.g., ultrasound). 4.7

C-15. SwRI - Chanute visit instructor's comment: There should be better selection and training opportunities for instructors. 4.1

C-16. SwRI - Chanute visit instructor's comment: Initial training instructors need more feedback from the field. 4.6

C-17. SwRI - Chanute visit: Courses are paced to slowest individuals; there should be more opportunity for advanced students to access additional information. 4.4

C-18. SwRI - Chanute visit: There should be more hands-on and practical tests. 3.1, 4.2

C-19. SwRI - Chanute visit: Instructors sometimes do not have extensive field experience. Chanute instructors suggest that 7-8 years experience should be required before a person can become instructor. 4.1

C-20. SwRI - Chanute visit: Instructors can become rigid and burn out. 4.1

C-21. SwRI - Chanute visit: Not enough visual training materials for a highly visual field. There are some, but they are infrequently used. 4.5

RESIDENT TRAINING

GENERAL NDE

D-1. Karimi: There is high rate of failure in EPRI's training course qualification exams. 4.3

D-2. Karimi: There is a large disparity between performance on qualification test and performance in the field. 4.9

D-3. Karimi: Basic training is operating under different set of values, standards, or directives from field environment. 4.10

D-4. Brock et al.: Report that training and selection are related to performance in radiography but amount of training was not related to performance in UT. 4.9

D-5. Bipes et al.: Initial training may be inadequate (but no data). 4.9

D-6. Burley: There is a need to improve training. 4.9

D-7. Zong: Improved training is the most important step that can be taken now. 4.9

D-8. Spanner et al.: Training is of utmost importance and hands-on training is best - currently there are no requirements for training upgrades. 4.2

D-9. Triggs et al.: They discuss whether learning is one of two types: detecting signal in noise or pattern recognition - they argue that we need to know which it is to determine implications for training. 4.8

D-10. Triggs et al.: Movement paths and orientation are important training/performance components. 4.9

ON-THE-JOB TRAINING

AIR FORCE NDI

E-1. SwRI observation: There is no computer-based trainer/tester that could help with Air Force sustainment training and impartial proficiency measurement during/after extension courses. 1.1, 5.1

E-2. SwRI - Laughlin visit: Air Force needs to address question of training of theory in the field. 5.1

E-3. ALC: There is sometimes a lack of up-to-date training material and information on new testing equipment delivered to the field. 5.7

E-4. Eagle, 1985: There is a need for "timely training of new techniques . . . upon arrival of new aircraft ...." 5.7

E-5. SwRI observation: OJT instructor selection and qualifications need to be standardized/re-assessed; the Air Force needs to standardize and exercise a quality-control program. 5.2

E-6. Summers: ". . . correct operation of the flaw detector instrument is a problem for new technicians and for those who use the instrument infrequently." The implication is that there is a need to provide refresher training for infrequent techniques. 5.8

E-7. Summers: Author felt motor manipulation of transducer was important enough component of OJT to warrant one of 4 modules on proposed computer-based NDI trainer. 5.1

E-8. SwRI - Laughlin Visit: There is a possible lack of training materials at OJT level. 5.1

E-9. SwRI - Laughlin Visit: Overall quality of OJT was rated high, but several indicated that it could vary as a function of the individual instructor. 5.2

E-10. SwRI - Laughlin Visit: There could sometimes be a double-bind situation in which the NDI technician's OJT instructor is also the person who rates his/her efficiency. An unwillingness to ask questions or admit not understanding something might result. 5.2

E-11. SwRI - Laughlin Visit: There was a reported lack of structure and standardization of OJT. 5.1

E-12. SwRI - Air Force interactions: There should be more emphasis on train-the-trainer for OJT instructors, and the system should insure high motivation among the trainers. 5.2

E-13. SwRI - Air Force interactions: More technicians should go back for advanced course, and instructors and shop chiefs also should go back for more training. 5.4, 5.5

E-14. Eagle, 1985: The problem is not the basic course at Chanute, rather, refresher training for widely scattered locations and OJT. 5.3

E-15. Eagle, 1985: There is a need for increased standardization of OJT for the -5 level technician. 5.2

E-16. SwRI - Laughlin visit: There is a diversity in OJT due to the fact that different installations use different NDI techniques. 5.3

E-17. SwRI - Chanute visit: The 7-level course is only optional and involves only a refresher of material covered in the initial training course. 5.5

E-18. SwRI - Laughlin visit: There is no independent critical assessment of hands-on performance or theoretical knowledge at critical points throughout the NDI career progression. 5.6

E-19. SwRI observation: There is a need for more feedback for training, performance, management, accountability, reliability, motivation, self-efficacy, meaningfulness, importance, justifying promotions, etc. 5.1

E-20. SwRI - Laughlin Visit: Some technicians felt that they were not kept up on current advances in NDI. 5.1

ON-THE-JOB TRAINING

GENERAL NDE

F-1. Rummel: OJT feedback is insufficient. 5.1

WORK ENVIRONMENT

GENERAL NDE

G-1. Wheeler et al.: Loud noise that is typical of the test environment interferes with communication. 6.1

G-2. Herr: The test environment disrupts performance: temperature, noise, light, etc. 6.1

G-3. Triggs et al.: Consider environmental variables; for example, Ramsey (1983) showed that heat stress can disrupt information processing performance. 6.1

G-4. Spanner et al.: Forsten and Aaltio; Peddada and Bennet report that time on task is important; also, Nakatsuji, Kuramochi, and Fujimori (1983) found that length of workday affected performance - later in the day yielded poorer performance. 6.1

G-5. Spanner et al.: Heat stress can cause performance decrement. 6.1

G-6. Spanner et al.: Physical space available determines performance. 6.1

G-7. Ainsworth: The literature on visual inspection shows that performance starts to degrade after about 40 min. without break. 6.1

G-8. Triggs et al.: State that no research has been carried out to assess effects of environmental variables on NDE performance but that related work hints there are effects. 6.2

G-9. Spanner et al.: Research is needed to determine how reliably measures obtained in a laboratory generalize to a hot, long, humid, heavy, and awkward real-world inspection tasks. 6.2

G-10. Wheeler et al.: Technicians are influenced by the physical "environment" (e.g., in immaculate plants, technicians get idea that high performance is demanded compared to dirty, unkept plants). 6.1

G-11. Wheeler et al.: Seven of twelve felt that working around schedules of others made it difficult to do good job. 6.1

POLICY

AIR FORCE NDI:

H-1. Summers: Technicians become specialized in certain techniques even though it may go against the "generalist" Air Force concept of NDI. There are possible disconnects between the general policy and the way that technicians are allowed to operate in actual field conditions. 1.3

H-2. Summers: The ASNT certified technicians performed better than their non-certified colleagues. 5.6

H-3. SwRI - Laughlin Visit: There was no evidence indicating that the technicians were not proficient as "operators," (following steps specified in a technical order). There was some evidence indicating that some of the NDI personnel might not have the theoretical knowledge necessary to adapt to a novel test situation or function in settings where there is no appropriate technical order. This issue is critical and could explain why NDI technicians have not always done well in some evaluations where flexibility was required or where they were required to perform inspections which were new or unfamiliar to them. This does not mean that the technicians cannot or are not performing their daily tasks at a high level (i.e., following technical orders to detect flaws), assuming that the technical orders are accurate, clearly presented, and cover all reasonable variations.

H-4. SwRI - Laughlin visit: Personnel reported that individuals who were gifted at certain techniques tended to specialize in those techniques. Apparently this is more common at some non-ATC laboratories where only a few techniques are commonly practiced. 1.3

POLICY

GENERAL NDE

I-1. SwRI summary: Numerous authors argue for stronger certification and policing of certification. 5.6

I-2. Wheeler et al.: Level III technicians are not significantly better than Level II despite greater certification demands. 5.6

I-3. Bush et al.: To prevent false optimism about NDE technician proficiency, "there should be a requirement" that the same equipment and procedures be used in the plant and technicians be required to pass demonstration tests. 5.6

I-4. Bush et al.: The experienced people won't like idea of being tested again and again and having to continuously go to school. 5.6

PERFORMANCE

AIR FORCE NDI

J-1. Summers: Inspections are infrequent; flaws are also infrequent. 7.1

J-2. Summers: There is a lack of augmented feedback; an automated trainer could provide this immediate feedback. 7.1

J-3. Lewis et al. (1978): Performance on assembled structures is poorer than that on components (Sproat, 1972). 7.3

J-4. Lewis et al. (1978): Human factors should be investigated. 7.2

J-5. Lewis et al. Persons were tested in this study who were not familiar with the test being conducted. 1.1, 1.2

J-6. SwRI - Chanute visit: One individual reported being tested in the "have cracks will travel" study on a procedure he/she had never conducted; despite his/her protests, the test was conducted and his data were included. 1.1, 1.2

J-7. Lewis et al.: Performance of NDI technicians falls below what has been assumed (MIL-A-83444 "Airplane Damage Tolerance Requirements" specifies 90% probability of detection with 95% confidence bound for all cracks .25 inches from the side of hole in thickness less than .25 inches). Results (except for florescent penetrant) indicated that technicians had trouble with 50-95. However, new equipment may be better (e.g., semi-automatic eddy and ultrasound - although entered too late in study - showed real promise at meeting guidelines). 7.4

J-8. Sproat: Fatigue crack detection probability is significantly lower than assumed. 7.4

J-9. Sproat: It is NDI proficiency and not equipment that accounts for the problem. 7.4

J-10. SwRI - Laughlin Visit: Technicians indicated that at Laughlin, they had the opportunity to practice all of the different Air Force NDI techniques, but that this was not always true of some other non-ATC assignments. 1.3, 7.5

PERFORMANCE

GENERAL NDE

K-1. Wheeler et al.: Technicians with greater experience and qualification under I&E Bulletin 83-02 (MRR - mini round robin) did not perform significantly better than other technicians (PIRR - pipe inspection round robin). 5.6

K-2. Wheeler et al.: "Far-side detection is not much better than pure chance." 7.4

K-3. Ainsworth: The performance of technicians has not kept up with capabilities of the equipment. In PISC I trials, technicians found less than 75% of 50 mm cracks using equipment capable of detecting cracks only a few mm long. 7.4

K-4. Bush et al.: This report states "... the UT examiner is one of the more erratic inspection variables ... " 7.4

K-5. Spanner et al.: Performance on tests has been disappointing and performance in field is unknown (but likely to be worse). 7.4

K-6. Triggs et al.: Existing performance data are test data; actual performance data remains unknown. 1.1

K-7. Wheeler et al.: There are problems to the extent that the following two lists of sources of information and consultation are not available:

Ranks for aids in <u>identifying</u> cracks:

1. Weld History
2. As-Built Drawings
3. Consultation with Same-Level Technician
4. Records of Inspection
5. Consultation with Higher-Level Technician
6. Consultation with Plant Engineering Staff

Ranks for aids in <u>sizing</u> cracks:

1. Consultation with Higher-Level Technician
2. Consultation with Same-Level Technician
3. Records of Inspection
4. As-Built Drawings
5. Weld History
6. Consultation with Plant Engineering Staff
   7.3

K-8. Herr: Prior manufacturing or maintenance sometimes have the effect of obscuring faults. (There are implications of this for the operator/TO approach.) 1.2, 7.3

K-9. Enrietto: The problem is too many false alarms, causing repairs that aren't needed and that actually can weaken system. 7.4

K-10. Rain: Adding requirements of location, size, depth and type to "detect" is what has increased human performance demands. 1.2, 7.6

K-11. Triggs et al.: Results from the related field of vigilance indicate that the following variables are important: task duration, task complexity, temporal distribution of signals, signal type, type of feedback, sleep loss, environmental stressors, work-rest cycles, and personality. 7.2

K-12. Triggs et al.: There is too little feedback, and the delay is too long. 7.1

K-13. Glasch: More research is needed to support human-factors. 7.2

K-14. Mordfin: The human factors of NDE are poorly understood. 7.2

K-15. Triggs et al.: Little systematic research has been conducted in human factors, and sometimes there have been confounded variables. 7.2

K-16. Fong: We may be reaching a limit on what any reasonably trained person can do. 1.2, 7.6

K-17. SwRI - Laughlin Trip: Personnel reported working in teams. No research could be found which investigated the effects of social or interpersonal relationships on NDI performance. 7.2

K-18. Ainsworth: In ultrasonics, little research has been conducted to isolate factors that could influence technician's performance. 7.2

K-19. Burley: Demands have been put on NDE that exceed state of the art. 7.6

K-20. Wheeler et al.: There is a lack of recent experience for Level III technicians. 7.5

K-21. Ainsworth: There is more attention paid to how equipment can be improved than how to improve human performance. He cites a review he conducted that found only 14 studies of human performance with UT. 7.2

PROCEDURES

AIR FORCE NDI

L-1. SwRI observation: There are no human reliability analyses to determine other sources and probabilities of error (besides the detection task) in the various NDI tasks. 8.1

L-2. SwRI observation: The more the shift to the operator mode, the more reliance is shifting toward perfect technical orders. Consequently, strict quality assurance procedures must insure that technical orders be carefully developed, accurate, de-bugged for misinterpretation, and periodically updated with information from the field. 1.2, 8.2

L-3. Lewis et al.: They report that performance using eddy current was better than performance using ultrasound on the same test piece. There is a potential concern if the individuals writing the technical orders do not take into account the skills and knowledge of the technicians using them.

L-4. Petru: Not only is the problem technician proficiency, but also, correct NDI engineering and effective process control are needed.

PROCEDURES

GENERAL NDE

M-1. Whittle and Coffey: Data-recording errors occur at surprising frequency, even in laboratory experience. 8.1

M-2. Wheeler et al.: Most technicians said that the layout of report forms was in wrong order and that this could lead to errors. The order is not same as the order of the information gathered. 8.2

M-3. Wheeler et al.: Technicians generally agreed that better detailed procedures are needed, new procedures don't always accompany new technology, and complying with approved procedures may yield faulty tests. 8.2

M-4. Herr: The common information provided to inspectors is that all parts are important and this is not true. Some are critical and also fracture prone. 8.2

M-5. Triggs et al. One possible solution to proficiency is to increase procedures; however, this could increase monotony and boredom. 8.2

M-6. Bush et al.: There is a need to continue to develop adequate criteria for qualification of entire inspection process. 1.1, 8.1

M-7. Rummel: Part of the problem is that the wrong NDE method might be selected; the examination not controlled; and NDE equipment, materials, and processes not controlled. 1.1, 8.1

M-8. Whittle and Coffey: Transcription errors have been found. 1.1, 8.1

M-9. Wheeler et al.: Twelve technicians' mean estimated error rate in recording data was 7%. 1.1, 8.1

M-10. SwRI observation: There are known instances in which standard procedures have systematically not been successful in finding defects. Only flexible technicians were able to identify the failure and make corrections. 8.2

MOTIVATION:

AIR FORCE NDI

N-1. SwRI observation: With no good measurement system, lack of feedback, and lack of practice in some methods, there is a corresponding lack of self confidence, accountability and corresponding reduction in motivation. 1.1, 9.1, 9.3, 9.5

N-2. Summers: Technicians were not confident in their ability. 9.3

N-3. Summers: There was no correlation between self-efficacy and performance; people don't know how good/bad they are. 1.1, 9.1, 9.3

N-4. SwRI - Laughlin visit: Some personnel feel that the maintenance community does not understand or appreciate the service that NDI provides. 9.4

N-5. Summers: There is evidence that motivation is important- people intending to reenlist in AF "seemed" to do better. However, it is difficult to determine which is cause and effect. 9.5

MOTIVATION

GENERAL NDE

O-1. Karimi: Technicians might have low motivation because don't have the necessary skills. 9.5

O-2. Karimi: May have low motivation because they are unskilled at making decisions. 9.5

O-3. Karimi: Informational and supportive feedback is not available in field, so technicians develop "negative beliefs" about their capabilities and opportunities. 9.1, 9.3

O-4. Wheeler et al.: There is a lack of confidence among technicians. 9.3

O-5. Wheeler et al.: There is a potential bias for technicians because ". . . crack calls will be rechecked by the utility, while non-crack calls will not be rechecked." 9.2

O-6. Spanner et al.: The technician's criterion can be influenced by subtle variables including management's unspoken attitudes. 9.2

O-7. Wheeler et al.: Some company supervisors apparently create biases in inspectors, while stressing the importance of finding flaws; 9 of 12 agreed that supervisors also stress the importance of being certain about the existence of any reported flaws. 9.2

O-8. Wheeler et al.: Technicians feel that utility's management is more concerned about false calls than correct calls - to a greater degree than their own company's supervisor. 9.2

O-9. Wheeler et al.: There is a lack of feedback ".. general dissatisfaction of technicians with the amount of feedback that they receive concerning their work." 9.1

O-10. Wheeler et al.: There is a general lack of feedback- only 3 of 12 said they received feedback more than 50% of the time; 5 of 12 said they received feedback less than 10% of the time. Yet, most indicated how important it was to get feedback. In addition: " . . . this lack of feedback gives them the impression that the utilities do not take their efforts seriously and does nothing to adjust their expectations (and thus their decision criteria) to the real state of affairs." 9.1

O-11. Wheeler et al.: Some companies do not stress importance of professional attitude in the technicians and the support staff. They give an example of technician who quit because clerks kept losing information, mistyping reports, etc. 9.4

O-12. Wheeler et al.: Technicians reported 2 kinds of fatigue problems: fatigue from long hours (more than 14) combined with other factors (e.g., wearing protective gear), and fatigue from working continuously over several weeks 5 or more days/week. They asserted that the former causes single incident errors and latter causes attitude/performance problems. 6.2, 9.3, 9.5

O-13. Rummel: Various motivational factors are implicated in NDE poor performance including: drowsiness, lack of interest, lack of motivation, fatigue, boredom, and monotony. 6.2, 9.3, 9.5

O-14. Kino: Nondestructive testing is boring. 8.2, 9.1, 9.5

O-15. SwRI observation: There is evidence of low motivation in a task that could be very motivating if the contingencies were right (i.e., looking for something that could yield great benefits if found, but is hard to find, by using skill and knowledge is what motivates most scientists, treasure hunters, and entrepreneurs). 9.5

O-16. Karimi: There is a large training - field disparity: a) cannot endure "physical hardships" of plants and b) still operating under different set of values, standards, or directives from field environment. 4.10, 9.5

O-17. Wheeler et al.: There is a lack of knowledge about procedure and lack of feedback following calls - "In most cases the technician did not know what the final determination was or how it was arrived at." 9.1

SYSTEM:

AIR FORCE NDI

P-1. Summers: Most frequent suggestion was more hands-on practice. 10.2

P-2. Summers: Supervisor's ratings of technician proficiency were not significantly correlated with performance. 10.3

P-3. Summers: There was wide variability in performance across technicians, bases, and commands in performance. 10.1

P-4. Summers: With regard to motivation of managers: ". . . inspection results were too inconsistent for maintenance managers to have confidence in NDI capability." 9.3, 10.3

P-5. Lewis et al.: There was heterogeneity in the types of techniques used at the different installations: un-weighted means over all test sites = 7.52% ultrasonic; 12.1% eddy current (surface); 5.62% eddy current (bolt hole); 50.19% penetrant; 24.57% X-ray. Also, within-technique ranges were broad: ultrasound 0% to 26%; eddy current - surface 1% to 27%; eddy current - bolt hole 0% to 16%; penetrant 4% to 90%; X-ray 1% to 74%. 10.1

P-6. Lewis et al.: There was heterogeneity in type of technicians: civilians and military; full time (e.g. Kelley-ultrasound) to part-time NDI (e.g., Kelley sheet-metal mechanics to Kelley radiographers), etc. 10.1

P-7. Lewis et al.: There was heterogeneity in tasks at different bases, a different mix of technicians (e.g., 100% civilian - Kelley to 100% military - several). Also, depots_ others on radiography; depot ¢ others on eddy current bolt-hole. 10.1

MEASUREMENT:

AF NDI:

Q-1. SwRI observation: In an area like NDI, lack of a good measurement system and the corresponding lack of feedback mean that non-performance or poor performance can go undetected. 1.1, 10.3, 11.5

Q-2. SwRI observation: To assure accountability, more valid performance measures must be introduced. 1.1, 10.3, 11.5

Q-3. SwRI observation: The existing test results are not probably true estimates of field proficiency. The technician being tested knows there is likely to be a fault in the sample; there are probably Rosenthal effects; motivation is probably abnormally high; there are probably Hawthorne effects; and there was no control or measure of response bias. 1.1, 11.3

Q-4. Boisvert et al.: Authors assert that "Have Cracks Will Travel" showed that pencil tests of theory did not predict proficiency. However, it is not clear that they have taken forgetting into account. The fact that test scores in initial training didn't correlate does not mean that later knowledge of theory will not correlate. 1.1, 11.3, 11.4

Q-5. Summers (personal communication): The proficiency problem may be exaggerated by the fact that NDI technicians tend to specialize in one technique but formal evaluations have tested them in several different methods. 1.1, 11.3

Q-6. SwRI comment on Lewis et al.: The bad grades may not be as bad as they would appear for training. For example, the "amount of training" comparisons are virtually meaningless because they don't distinguish between <u>what</u> techniques were trained. There is also a chance that only poor technicians might be selected for additional training. If so, results are showing that training works, not that training doesn't work. 1.1, 11.3, 11.4

Q-7. SwRI comment on Lewis et al.: The absence of statistics in the sections of the report making comparisons among different variables make it difficult to assess the outcomes. 1.1, 11.3

Q-8. SwRI comment on Lewis et al.: It is also apparent that the lack of relationship of experience and proficiency is based on total years of experience, not based on experience on the technique being measured. To support the conclusions drawn, analyses are needed demonstrating correlations between hours of eddy training and eddy-current performance, hours of ultrasonic practice and ultrasonic performance, etc. 1.1, 11.3, 11.4

Q-9. SwRI comment on Lewis et al.: Authors stated that there is "not yet a generally agreed upon method for analyzing the impact of these false calls on NDI reliability." There is no attempt to deal with false calls. In fact, they say that there were some extremely high false-call rates in the radiography tests. 11.1, 11.2

Q-10. SwRI comment on Lewis et al.: It is difficult to determine what quality control was used for the equipment, to insure that technician, and not equipment, proficiency was measured. 1.1, 11.3

Q-11. SwRI - Chanute visit: One individual reported that during participation in "have cracks will travel" he/she was asked to conduct a test never-before conducted by that individual. This individual felt that the results were biased from such an approach and that he/she was really a proficient technician when conducting familiar tests. 1.1, 11.3

Q-12. Sproat: He noted the problem of the older measurement technique of obtaining 95% confidence interval (too conservative) and referenced Berens & Hovey (1981) who provided a better mathematical approach. 11.1

Q-13. Boisvert et al.: Existing personnel qualifications documents do not measure a technician's proficiency. MIL-STD-410 and SNT-TC-1A test theoretical knowledge and ability to operate equipment, but no real practical test is specified. 1.1, 5.6, 11.3

Q-14. Boisvert et al.: Because of the large variety of tests and applications among laboratories, successful measurement at one facility may not mean success at another facility. A standardized measurement system is needed. 10.1, 11.3

Q-15. Lewis et al.: There is a need for a method to evaluate performance, skill development and motivation. 11.1, 11.2

Q-16. Swets (1983a and 1983b): The older probability of detection (POD) measures are not as good as signal detection analysis which can derive independent measures of detection (sensitivity) and response bias. 1.1, 11.1, 11.2

Q-17. Summers: There was no correlation between performance and like/dislike for present job or NDI career field. 11.1

Q-18. Summers: Time spent on tasks was not correlated with performance. 11.1

MEASUREMENT

GENERAL NDE:

R-1. Wheeler et al.: Authors state that field-like testing difficulties like weld reinforcement and ripple were not included in the test. Similarly the "weld reference system" which has difficulties was not used in tests. Also, there was no time limit in test, but there often are time limits in real field environment. The general theme is that real performance is probably worse than tests would make it seem. 1.1, 11.3

R-2. Glasch: The older POD measures are not as good as signal-detection measures. 1.1, 11.1, 11.2

R-3. Glasch: One serious shortcoming of the Air Force report was that it failed to report false alarms (hence, no measure of sensitivity or bias can be computed). 1.1, 11.3

R-4. Spanner et al.: Some measures such as POD (probability of detection), sensitivity, and specificity have flaws; ROC is recommended because those flaws are removed. 1.1, 11.1, 11.2

R-5. Triggs et al.: Most of the traditional measures of performance are not adequate. 1.1, 11.1, 11.2

R-6. Mordfin: "There is a need for a formal and widely accepted methodology for assessing NDE reliability . . ." 1.1, 11.1

R-7. Mordfin: There is no formal methodology for assessing NDE reliability. 1.1, 11.3

R-8. Spanner et al.: Of the three traditional measures of detection (POD, sensitivity, and specificity), none is acceptable. 1.1, 11.1

R-9. Triggs et al.: In creating test samples, great care should be taken not to over-screen candidates, creating a sample that is not representative. 11.3

R-10. Triggs et al.: Describe work of Lusted et al. (1977) dealing with assessment of technique/system effectiveness that addresses three levels of efficacy. Only first level (diagnostic-thinking efficacy) is typically addressed in NDE proficiency measures. 1.1, 8.1, 11.3

R-11. Ainsworth: (Not from NDI literature but from visual inspection literature) The "fault probability effect" states that as proportion of faulty material decreases, the probability of rejecting faulty work is reduced. Consequently, test results of laboratory tests are probably an underestimate of technicians' proficiency. 11.2

R-12. Herr: Accept/reject criteria are often vague. 11.2

R-13. Wheeler et al.: Some of the tests lack construct validity, ". . . technique for passing the NDE Center performance demonstration is unrelated to the manner in which decisions are made in the field." 8.1, 11.3

R-14. Wheeler et al.: They stated there was apparently more concern with checking calls than no-calls in real life. 11.2

R-15. Bush et al.: Technicians are susceptible to existing engineering expectations. "Before .. an operator claiming to have found a crack ... could expect to meet skeptical resistance, because of the general perception that these pipes were not susceptible to IGSCC." 11.2

R-16. Swets (1977): Supervisors can create more response bias against false calls than missed calls. 11.2

R-17. Triggs et al.: More supervision skills are required because of subtle nature of work. In addition, false calls are more likely to be caught by supervisors than missed calls. 11.2

R-18. Triggs et al.: Not much research has been conducted in the area of procedural variables in NDE. 1.1, 8.1, 11.3

R-19. Mordfin: NDE system reliability is currently uncharacterized. 1.1, 8.1, 11.3

R-20. Fong: One cannot characterize NDE system until a systems approach is taken. 1.1, 8.1, 11.3

R-21. SwRI from Triggs et al.: Of the 10 steps described in the generic NDT procedure, most attention is paid to the "material examination" step, possibly because it is most salient and face valid test of proficiency. The other 9 are testable but have not been measured to the same extent. 1.1, 8.1, 11.3

EQUIPMENT:

AF NDI:

S-1. Lewis et al.: The semi-automatic ultrasound and automatic eddy current techniques were better than their non-automated counterparts; therefore, the equipment could be more automated. 12.1

S-2. Lewis et al.: Equipment effects were found in this study. 12.2

S-3. SwRI - Chanute visit: NDI technicians do not always have the same equipment for initial training that they will be using in the field. 12.2

EQUIPMENT

GENERAL NDE:

T-1. Herr: There is no standardization of names, functions, knobs, and controls of testing equipment. 12.2

T-2. Busse, et al.; Posakony: There is a wide variation in equipment and, worse, most technicians are not aware of the wide range in equipment variations and how they can affect overall performance. 12.2

T-3. Burley: Technicians are asked to use instruments with different sensitivities, standards of unknown traceability, and techniques that may or may not have been standardized for particular material or structure. 12.2

T-4. Wheeler et al.: "Hard to manipulate controls, which are difficult to use when wearing protective gloves and are subject to accidental actuation." 12.3

T-5. Wheeler et al.: "Search units that are difficult to hold and control due to their small size and lack of grip aids." 12.3

T-6. Herr and Marsh: Equipment controls are not standardized. 12.2

T-7. Triggs et al.: Equipment controls and displays are not standardized. 12.2

T-8. Wheeler et al.: "Displays, labels, and scales that are difficult to read at the viewing distances that may be encountered during ISI/IGSCC inspections." 12.3

T-9. Wheeler et al.: The technicians wanted calibrated time base, clear displays, and used location instead of amplitude to detect faults. 12.3

T-10. Spanner et al.: The layout, labeling, signal processing, and screen clarity of images need to be evaluated. 12.3

T-11. SwRI comment: With new NDI equipment procurements and the high reliance on technical orders, great care must be made to update technical orders to correspond to new equipment. 8.2, 12.4

T-12. Spanner et al: Equipment is sometimes difficult to read, even in perfect lighting, also sometimes protective masks or plastic covers to reduce radiation make it difficult to read display. 12.3

T-13. Watkins and Cowburn: UT displays are not effective. 12.3

T-14. Spanner et al.: Effectiveness of information display has been questioned (Watkins and Cowburn, 1980). 12.3

T-15. SwRI comment from Wheeler et al.: The broad range of equipment ratings (poor to excellent) suggests individual differences that are being accommodated. There are no data to delineate these or decide which should be adjustable or constant. 12.3

T-16. Herr: Equipment manufacturers keep producing equipment that is not evaluated for human's capability to use. 12.3

T-17. Herr: Transducers vary greatly and some are bad out of the box. 12.3

T-18. Triggs et al: Shape or color coding not used to differentiate search units. Labels are often etched on where gripping occurs. 12.3

T-19. Triggs et al.: In UT, return amplitude depends on: force on search unit, uniformity of pressure, surface roughness, type of couplant, and couplant thickness. Yet, little human factors work has been conducted on the probe except for Herr and Marsh, 1987. 12.3

T-20. Herr: Little information or help is available from manufacturers. 12.6

T-21. Kino: We need better equipment and more automated equipment (e.g., better imaging screen). 12.1, 12.3

T-22. Spanner et al.: The calibration can be affected by changes in search unit, connecting cable, etc. 8.2, 12.1

T-23. Jamison and McDearman: Equipment band-width of ultrasonic amplifiers can influence estimated fault size. 8.2, 12.1

T-24. Spanner et al.: Eitzen et al. (1975) showed that standard reference blocks can give different results. 8.2, 12.1

T-25. SwRI - from Triggs et al.: According to a task analysis, visual inspection is required simultaneously for keeping the sensor on the correct path and monitoring screen; these are concurrent tasks that are impossible to conduct simultaneously. 12.3, 12.5

T-26. Bush et al.: Currently, some pipe welds are inaccessible or very difficult to access for UT exam. 12.6

T-27. Bush et al.: Original design could be improved-currently, conditions exist that make NDI job difficult-geometry problems - counterbores, inner-surface cladding, mismatch, weld root drop-through, suck-back - these can cause false alarms or affect criterion, hence failure to call defect. 12.6

T-28. Kino: The reliability of NDI transducers is very low. 12.2

CONCERNS NOT USED BECAUSE THEY WERE JUDGED INAPPROPRIATE:

1. Herr: argued that selection is important to NDE, but limits set by Federal law and union agreements restrict selection process.

2. Lautzenheiser et al.: Discusses the need for more training in schools.

3. Wheeler et al.: They had surprising finding that near-side short defects were easier than near-side long defects. They said it might have been an artifact due to the grading unit. They corrected, redid the analysis, and found that the difference was in sizing not in detecting long versus short. Nevertheless, no difference in long/short near side is counter-intuitive; it seems long would be easier. In addition, the long cracks are obviously more hazardous.

4. Ainsworth: Several studies have shown that large cracks are found more often than small, that small cracks are overestimated in size, and that large cracks are underestimated in size (this last finding can lead to dangerous situations).

5. Bush et al.: "While length sizing of cracks is acceptable, depth sizing is currently inadequate."

6. Bush et al.: There is currently no good technique for detecting and dimensioning flaws in pipes repaired by the weld overlay process.

7. Wheeler et al.: They presented significant correlations showing that the more experience a technician has the <u>worse</u> at detecting easy cracks (near-side short) and the better at detecting hard cracks (far-side long).

8. Bush et al.: Because of grain boundaries, applying conventional shear wave exam required by Code, . . . the far side heat-affected zone (HAZ) may not be inspected.

9. Bush et al.: Results of EPRI report (EPRI/NDEC 82-10-Aug1983) indicated a tendency to oversize cracks less than 20% through wall and to underestimate those deeper than 20%. The number of teams with adequate performance was less than anticipated. There was a huge range in performance. The crack-tip diffraction sizing approach was found to be the most viable method. Corrective actions are needed. None of the teams could size accurately. Most field cracks have been found to be tight; therefore, basing sizing performance estimates on EDM notches and wide cracks is "non-conservative."

10.. Lewis et al.: Temperature differences found were opposite to those expected if due to temperature effects on human, therefore, must be due to technique.

11. Wheeler et al.: Protective clothing and gloves make it hard to move around.

12. Wheeler et al.: Face masks become fogged with heat and humidity.

13. Bush et al.: Currently, testing often occurs under very imperfect conditions: high temperature and humidity, respiration equipment, protective clothing, awkward positions, time limits, etc.

14. Summers: Performance was not related to degree of comfort/discomfort with the equipment used.

15. Ainsworth: Performance in radiography is dependent on film type and "nature of the viewing screen."

16. Spanner: Different teams using different procedures make different calls on same welds.

17. Spanner et al.: Most of the UT/ISI procedures are so generalized that wide latitude is permitted.

18. Forsten and Aaltio: There is incorrect reporting because of erroneous coordinates.

19. Bush et al.: Improved procedures are needed - 60-degree shear wave added to 45-degree shear wave for increased detection.

20. Bush et al.: Improved procedures are needed - skewed scan added to current parallel and perpendicular scans.

21. Spanner et al.: ASME Code rules don't provide guidance for interpreting complex signal patterns that are found in welds.

22. Karimi: There is a lack of goals or conflict between goals.

23. Wheeler et al.: Both self-efficacy and performance were lower for far-side cracks.

24. Wheeler: ". . . different technicians scanning the same pipe will produce different results," ". . . formal certification level is not related to UT ability."

25. Wheeler et al.: Some technicians have a fear of radiation and an adverse reaction to the pressure of time limits due to radiation.

26. Kemmler: Described survey covering all organizational levels showing that there is no strong interest in quality improvement unless prescribed by code or law or purchaser.

27. Spanner: The stress/performance tradeoff is problem.

28. Doctor: There is little improvement due to IEB 83-02 requirements.

29. Wheeler et al.: Usually two techniques are used in UT/ISI (inservice inspection). One is where the same technician scans the pipe and watches the display; the other is a master/slave (M/S) arrangement where one technician scans and other watches the display. What are performance implications for these? The technicians reported problems with M/S - the technician doing the scan must go at a constant scrub rate, must have confidence in technician doing scrub, calls are not made without individual inspection. One gets the idea that M/S is never sufficient. If M/S is widely used, research is needed to compare effectiveness.

30. Bush et al.: The original design could be improved - grain size variations in pipe, heat affected zone (HAZ), different metal in weld can attenuate beam and return irrelevant signals; these are worst in stainless steel material and welds.

31. Hellier (1980, 1984): The NDE certification by individual companies, as specified by SNT-TC-1A, is not working, gives examples of violations.

32. Zong: The qualification and certification program is a myth.

33. Hellier : "Many of the <u>practical</u> examinations administered are a futile exercise."

34. Rummel: There is a problem in NDE management, training and materials control.

35. Mordfin: Some of the newer NDE methods are inadequately standardized.

36. Fong: There is not enough technical basis in codes and standards.

37. Burley: Part of problem is economic: management sees NDE as overhead, others as dangerous to the product.

38. Spanner et al.: Technician is supposed to study previous weld data before exam, but quantity of useful information varies from plant to plant.

39. Triggs et al.: In referencing location data, communication to the assistant is required - but often difficult in that environment.

40. Wheeler et al.: Technicians said that batteries are poorly maintained.

41. Wheeler et al.: Equipment is currently not suited to hostile radioactive environment.

42. Triggs et al.: Reducing the probe frequency decreases the effects of orientation (Taylor and Selby, 1981) but at cost of degrading the signal (Haines, 1977).

43. Herr: Perishable equipment is not satisfactorily controlled.

44. Bush et al.: Calibration blocks should be improved (e.g., block = pipe material, welds included, etc. p 4-6).

45. Triggs et al.: Need to know more about what effect beam profile and frequency have on performance (Watkins and Cowburn, 1980).

46. Chin Quan: There is controversy about whether increased gain and other significant controls improve performance. Chin Quan assumed increase but found opposite.

47. Ainsworth: There is a dilemma in the adjustment of sensitivity: increase in sensitivity also increases noise.

B. Summarized Areas of Concern and Criticality Ratings

To reduce redundancy among originally identified concerns, condensed concerns were created, intended to encapsulate all identified issues in a smaller and more manageable set. The condensed concerns (along with all the reinforcing initial concerns) were presented to the NDI program office for their review and comment.

In the course of this exercise, it became apparent that there were three significant "overriding" issues that impacted a substantial set of the concerns. One overriding issue dealt with the fact that there is currently no adequate measure of NDI performance. This issue has far-ranging affects on training, feedback, motivation, morale, and assessment. A second overriding issue dealt with one goal of the NDI technician program - was the goal to create a "technician" or an "operator?" The third overriding issue dealt with another goal of the NDI program - was the goal to create a generalist who was adequate at all possible tests, or a specialist, who was expert in a smaller number of tests?

Because the SwRI team did not have the answer to this Air Force policy decision, two sets of "criticality" ratings were made for each condensed area of concern. In the margin to the left of the concern, the resulting mean criticality ratings are provided: the top value represents the judged criticality if an NDI operator is the desired goal, and the bottom value represents the judged criticality if an NDI technician is the desired goal. In the following pages, condensed areas of concern are presented, along with their supporting basic concerns (the alphanumeric codes following the stated concern), and their two criticality ratings.

1.0 Global Areas of Concern:

The areas of concern expressed in this section are of sufficient importance that they have been presented separately. In all cases, they transcend and impact specific areas of concern presented in the subsequent sections. We believe that these concerns must be resolved before any systematic program can hope to improve NDI proficiency.

1.1 The <u>Measurement</u> of NDI proficiency. J-5, J-6, K-6,
5.16 L-1, M-1, M-6, M-7, M-8, M-9, N-1, N-3, Q-1 to Q-8,
6.68 Q-10, Q-11, Q-13, R-1 to R-8, R-18 to R-21

Explanation: The concern is that there are no meaningful measures of NDI proficiency. Recent investigation of the measurement techniques used in many reports in which NDI proficiency was presumably assessed have shown that the measures used are inadequate. In addition to measurement per se, there are other questions about the tests conducted that raise questions about derived estimates of "true" NDI proficiency. Nevertheless, reports of low proficiency from those studies can probably serve as an index of proficiency (by our estimation, they are overestimates of true NDI <u>technician</u> proficiency). Because of the lack of a reliable and valid measurement of NDI proficiency, any concerns or solutions expressed in this report must be qualified as speculation, based on expert opinions and rather few empirical findings. The measurement issue has far-reaching implications for human performance including accurate feedback, training, accountability, motivation, and performance assessment and must be resolved. The fact is that we were currently unable to assess NDI proficiency.

1.2 The <u>Role</u> of Air Force NDI personnel: operator or
4.88 technician? H-3, J-5, J-6, K-8, K-16, L-2
6.65

Explanation: The concern is that the Air Force has not decided the extent of the skills that NDI personnel should have. To dramatize the point, a distinction is made between an "operator" and a "technician." The issue here goes far beyond semantics. An NDI "operator" is an individual who is familiar with NDI equipment and procedures, can follow appropriate detailed technical orders, and can come to a meaningful test outcome (i.e., flawed or not flawed). A technician is an individual who can do all that an operator can do, but in addition, because of a sound

understanding of the underlying theory, equipment, and methodologies, is flexible enough to adapt tests when technical orders are inappropriate or absent. The specification of the role of NDI personnel has implications for the selection of personnel, resident and on-the-job training, equipment features, and development of technical orders.

1.3 The Responsibility of Air Force NDI personnel: general
5.24 responsibility for all NDI techniques or specialist in
6.29 certain techniques. H-1, H-4, J-10

Explanation: The concern is that Air Force NDI personnel are currently expected to be proficient in many different NDI techniques. This concern is related to 1.2 above. Specifically, it is feasible to field "generalist" operators, but less likely that proficient "generalist" technicians could be selected/trained. For example, in private industry a technician claiming proficiency in more than 3-4 techniques would probably be questioned, because of the technical complexity involved in most techniques. The question is whether the Air Force should seek expert technicians, specializing in a few techniques, or expert operators, able to follow technical orders using a variety of techniques. The question is similar to that recently confronted for Air Force pilots, for which the solution was specialization.

Specific Areas of Concern

2.0 Selection

2.1 In a highly technical area, there is no intentional
3.67 selection mechanism. NDI technicians come from the
6.33 general manpower pool. Quality of performance could
probably be enhanced by selecting from the electronics
or mechanics personnel pools. A-1, A-4, A-5, B-1

2.2 There are a number of candidate selection variables
3.33 proposed in the technical literature, but little
6.67 systematic research into which variables predict good
NDI technicians. The literature that is available is
not meaningful since measures of the predicted variable
(proficiency) were not adequate. A-2, A-3, A-6, A-7,
B-2, B-3, B-4, B-5, B-8

2.3 Many of the tests involve visual inspection, yet there
3.92 are no specialized eye examinations given to help
5.31 select/confirm individuals for visual skills. B-6, B-7

3.0 Meta-training.

3.1 The general concern that transcends basic and on-the-
3.82 job training (OJT) is the distribution of training over
6.41 time. If the goal of NDI training is to produce a
proficient "technician," then formal training in theory
is too general and lacking in depth. If the goal is an
NDI "operator" then the current initial training is
probably adequate (if not too intensive) given the time
allowed. The goal must be specified before the
training can be tailored to accomplish that goal. C-1,
C-10, C-18

4.0 Initial Training

4.1 The selection/qualification system for NDI instructors
4.74 needs to be specified/revised. For example, some
6.00 instructors have no field experience. C-2, C-15, C-19,
C-20

4.2 Depending on the goal of NDI training, the amount and
5.33 mix of theory/hands-on experience in the basic course
6.33 need to be adjusted. C-10, C-18, D-8

4.3 The lack of attrition within the basic course is in
3.00 sharp contrast to the relatively high failure rate
6.38 found in many commercial training programs. D-1

4.4 The group-paced mode offers little freedom for advanced
4.00 or motivated individuals to elaborate their
6.67 understanding of the information. C-9, C-13, C-17

4.5 In a highly "visuospatial" field, there are relatively
4.67 few visual aids to facilitate understanding
5.67 (videotapes, films, computer animation, etc.). C-9,
C-12, C-21

4.6 There is little feedback to initial training
3.94 instructors from the field. C-16
6.00

4.7 The hours of training per technique should be re-
4.80 assessed. For example, the spectrometric oil analysis
6.47 program (SOAP) technique (which is highly automated and
sometimes not considered to be a NDE procedure)
receives 61 hours while ultrasound (a highly complex
and theoretical NDE procedure) only receives 51 hours.
C-14

4.8 Research is lacking, and findings in the current
4.31 literature have not been assessed to determine what
5.77 learning components are most important (e.g., theory,
motor-behavior, procedural training, pattern-
recognition, signal detection in noise, etc.) and which
training techniques are most successful (e.g., cuing,
knowledge of results, immediate feedback, etc.).
Findings are not used to actively update the
content/delivery of information in the basic course.
C-5, C-11, D-9

4.9 There is little evidence supporting the notion that
4.36 training improves NDI performance. C-3, C-4, C-7,
6.07 C-8, D-2, D-4, D-5, D-6, D-7, D-10

4.10 There is too great a discrepancy between initial
3.75 training and subsequent field environment. D-3, O-16
6.00

5.0 On-the-Job Training (OJT)

5.1 There is a lack of standardization, structure, and
6.06 training materials in the current OJT. E-1, E-2,
6.17 E-7, E-8, E-11, E-19, E-20, F-1

5.2 The selection, training, qualifications, and motivation
4.94 of the OJT instructor along with his/her organizational
6.44 relationship with the student need to be
standardized/re-assessed. E-5, E-9, E-10, E-12, E-15

5.3 There is a lack of diversity in OJT in some instances
4.81 due to the fact that some installations specialize in
5.44 certain NDI techniques because of the aircraft they
service. E-14, E-16

5.4 There is a lack of formal training from the completion
4.19 of initial training until the optional 7-level course
6.38 (and that is refresher training). E-13

5.5 The 7-level course is optional and does not introduce
3.56 added depth of information. E-13, E-17
6.69

5.6 There is no independent critical assessment of NDI
5.67 proficiency (hands-on or theoretical knowledge) at
6.33 critical milestones throughout the NDI career
progression. E-18, H-2, I-1, I-2, I-3, I-4, K-1
Q-13

5.7 There is occasionally a lack of update training
4.47 material and information on new testing equipment
4.80 delivered to the field. E-3, E-4

5.8 There is no vehicle to insure that personnel get
5.20 adequate practice using infrequent techniques. E-6
6.07

6.0 Work Environment

6.1 NDI technicians and their supervisors should be made
5.00 aware of the effects of environmental variables like
4.62 temperature, noise, amount of light, humidity,
extended time-on-task, and work schedule on human
performance. G-1 to G-7, G-10, G-11

6.2 There is a lack of research on the effects of work and
4.31 environmental variables on NDI performance. G-8, G-9
4.54

7.0 Performance

7.1 Because of the nature of NDI tasks, there is an
5.22 intrinsic lack of performance feedback. For example,
6.72 there is obviously no real-world immediate feedback for
missed calls, and often delayed or no feedback for hits
and false alarms. Implications extend to assessment,
training, motivation, accountability, and self-
confidence. J-1, J-2, K-12

7.2 There is a lack of systematic research in the area of
5.17 NDI performance. Also, what is known from other fields
5.69 is not often applied or communicated to technicians or
managers. J-4, K-11, K-13, K-14, K-15, K-17, K-18,
K-21

7.3 NDI tasks and procedures are often not designed with
6.47 human performance in mind. J-3, K-7, K-8
4.60

7.4 Nondestructive testing performance is not adequate.
5.35 (Virtually all authors agreed, but some made cogent
6.53 points: J-7 to J-9, K-2 to K-5, K-9.)

7.5 NDI personnel report barriers to practicing some
3.75 techniques at some installations. J-10, K-20
5.38

7.6 If the operator approach is taken, then temptations
6.43 must be resisted to increase demands on the operators.
3.57 K-10, K-16, K-19

8.0 Procedures

8.1 Most assessments of NDI proficiency have focussed on
3.33 the detection task. While this is the most important
5.67 task, there are other human tasks involved. Human
error can occur in any task and reduce proficiency.
No human reliability assessments could be found in
which error rates are assessed for all tasks/subtasks.
Consequently, even if the existing test results were
valid and were perfect indicators of actual field
performance, there would still be no true assessment of
actual NDI proficiency. L-1, M-1, M-6 to M9, R-10,
R-13, R-18 to R-21

8.2 To the extent that the Air Force embraces the operator
7.00 as opposed to the technician approach to NDI, there are
4.13 implications for the procedures used. Extensive
quality control procedures for technical orders must be
strictly followed and tested to insure that there is no
possibility of misinterpretation. In addition, steps
should be taken to minimize possible negative side-
effects such as boredom. L-2, L-3, L-4, M-2, M-3, M-4,
M-5, M-6, M-10, T-11, T-22 to T-24

9.0 Motivation

9.1 The lack of feedback that is characteristic of NDI
5.53 tasks affects the motivation of the NDI personnel.
6.41 N-1, N-3, O-3, O-9, O-10, O-17

9.2 The effects of response bias and the variables that
5.83 influence response bias have been ignored in most
6.25 research on NDI performance proficiency. O-5, O-6
O-7, O-8

9.3 There is a reported lack of confidence and self-
4.73 efficacy among nondestructive testing personnel.
6.00 N-1, N-2, N-3, O-3, O-4, P-4

9.4 The professionalism, feelings of importance in the
4.73 maintenance community, and general esprit de corps
6.33 could be improved for NDI. N-4, O-11

9.5 There are reports of low motivation among technicians
4.40 in the nondestructive testing community. N-1, N-5,
6.07 O-1, O-2, O-12 to O-16

10.0 System

10.1 There is a heterogeneity among personnel and
5.08 NDI tasks across different units, which makes it more
5.14 difficult to prescribe general, standardized remedies.
P-3, P-5 to P-7, Q-14

10.2 There needs to be more opportunity for practice,
5.39 especially for infrequently used procedures. P-1
6.00

10.3 Because of the lack of adequate measures of
4.40 performance, standards and qualifications for career
5.73 advancement are less defendable. P-2, P-4, Q-1, Q-2

11.0 Measurement

11.1 The measures of proficiency used in the past by
5.63 researchers investigating NDI technician performance
6.44 are not good measures of true detection ability
(sensitivity). Q-9, Q-12, Q-15 to Q-18, R-2, R-4 to
R-6, R-8

11.2 The measures of proficiency used in the past by
5.83 researchers investigating NDI technician performance
6.42 are not good measures of the technician's response
bias. Q-9, Q-15, Q-16, R-2, R-4, R-5, R-11, R-12,
R-14 to R-17

11.3 The experimental design or approach taken in the past
5.83 by researchers investigating NDI technician performance
6.42 were sometimes inappropriate or led to
misinterpretation. Q-3 to Q-8, Q-10, Q-11, Q-13, Q-14,
R-1, R-3, R-7, R-9, R-13, R-18 to R-21

11.4 Because of inappropriate measures of proficiency, there
5.83 might be inaccuracies in the literature with regard to
6.42 potential correlates of NDI proficiency. Q-4, Q-6, Q-8

11.5 The lack of an appropriate measure of proficiency has
5.83 potential wide-ranging negative effects ranging from
6.67 difficulty in training to low motivation/morale. Q-1, Q-2

12.0 Equipment

12.1 Equipment should be selected which is semi-automated
6.38 and reduces the number and complexity of the decisions
4.00 made by the inspector (especially if the NDI operator path is taken). S-1, T-21, T-23, T-24

12.2 NDI equipment (displays, controls, and sensors) should
6.14 be standardized as much as possible (especially if the
4.29 NDI operator path is taken). S-2, S-3, T-1 to T-3, T-6, T-7, T-28

12.3 NDI equipment displays, controls, and sensors should be
6.14 designed, evaluated, and selected on the basis of their
4.29 ease of use among actual NDI technicians. T-4, T-5, T-8 to T-10, T-12 to T-19, T-21

12.4 Because of the criticality of the technical orders in
6.14 Air Force NDI, great care must be taken to insure that
4.29 all technical orders correspond to current equipment available. T-11

12.5 There is a lack of knowledge and research in the areas
5.69 of display design, controls design, and sensor design
4.08 to maximize human performance. T-25

12.6 Weapon-system manufacturers do not always consider NDI
5.69 testing when designing a new system; NDI equipment
4.46 manufacturers are not always accessible. T-20, T-26, T-27

IV. RESULTS OF PHASE II: POSSIBLE SOLUTIONS

Possible solutions are presented in this section. To assist the reader, the global and specific areas of concern are first restated, in order that the identified possible solutions can be interpreted in terms of the concern they addressed. Each possible solution is identified by an alphanumeric code; however, small-case letters have been used to distinguish the possible solutions from the earlier concerns. Following the code, a brief reference is cited; more detailed reference information can be found in the References section.

1.0 Global Areas of Concern

1.1 The Measurement of NDI proficiency.

a-1-1. SwRI based on numerous authors: The Air Force should incorporate a standardized performance measurement system. The adopted system should provide both measures of proficiency (sensitivity) and response bias (the technician's overall tendency to make calls). The best candidate measures are based on signal detection theory, in which the two measures can be independently estimated. In addition, improved regular standardized tests of proficiency should be implemented.

1.2 The Role of Air Force NDI personnel: operator or technician?

and

1.3 The Responsibility of Air Force NDI personnel: general responsibility for all NDI techniques or specialist in certain techniques.

a-2-1 and a-3-1. Because these two issues are so highly related, they are treated jointly. The Air Force's personnel goals along these two dimensions should be formally defined so that training, expectations, tests, etc., can be adjusted accordingly. Theoretically, these two dimensions can be considered independent. Consider the table below:

		Responsibility	
		General (many techniques)	Specific (few techniques)
Role	Operator	A	B
	Technician	C	D

It is our impression from the literature and from site visits that, in general, technicians are more likely to fall into cell A (if at certain installations where many techniques are practiced) or cell B (if at certain installations where few techniques are practiced). There is nothing wrong with cells A and B, as long as certain quality assurance steps are taken and as long as the Air Force realizes the implications (e.g., the high reliance on technical orders).

It is unrealistic to set cell "C" as a target because a) experts in the field question whether cell "C" is attainable, and b) if it were, it would not be cost-effective to spend that much time and money in training individuals who would then be very attractive to outside employers. The Air Force should formally adopt a position; a compromise is recommended.

First, initial training and early on-the-job training (OJT) should have the generalist operator as a goal. Second, to conduct/monitor training, to conduct difficult or unusual tests, to monitor quality-control, and to make final decisions on controversial calls, there should be an appropriate number of specialist technicians at each installation. These individuals would be experts who have demonstrated high competence in certain techniques.

The -7 level course should become mandatory, selective, more advanced, and specialized to certain techniques. Graduates would serve as the specialist technicians discussed above. Tests of proficiency must be periodically conducted, but they must be limited to the targeted role of the individual.

This approach would be cost-performance effective because a) it is closest to what the Air Force is currently practicing b) minimizes the training time/expense required for the majority of NDI personnel while keeping proficiency at a reasonable level, and c) specifies realistic, obtainable, <u>testable</u> goals.

Specific Areas of Concern and Possible Solutions

2.0 Selection

2.1 In a highly technical area, there is no intentional selection mechanism. NDI technicians come from the general manpower pool. Quality of performance could probably be enhanced by selecting from the electronics or mechanics personnel pools.

b-1-1. Sproat: Found that AF NDI technicians could be highly proficient, if selectivity is exercised in personnel assignments.

b-1-2. Sproat: Identify individuals high and low in proficiency in certain tests (however, this goes against "universal NDI" technician policy).

b-1-3. SwRI comment on a-2: A fielded computer-based trainer that could also test proficiency could assist supervisors assess technician proficiency.

b-1-4. Ainsworth: Selection should be reconsidered with NDE in mind.

b-1-5. SwRI: The Air Force should test the hypothesis that NDI technicians should be drawn from the electronic or mechanical manpower pools instead of the general manpower pool. Such a test could be easily and cost-effectively conducted.

2.2 There are a number of candidate selection variables proposed in the technical literature, but little systematic research into which variables predict good NDI technicians. The literature that is available is not meaningful since measures of the predicted variable (proficiency) were not adequate.

b-2-1. Lewis et al.: Develop personnel screening program to select by intelligence, temperament, and dexterity.

b-2-2. Lewis et al.: Select on basis of proficiency and motivation.

b-2-3. SwRI: To provide appropriate judgments about what variables predict proficiency, the Lewis et al. data should be re-analyzed using ROC's correlations re-calculated using d' or other sensitivity measure. However, this is impossible due to the fact that false alarm data were not maintained. Consequently, any new data collected should include both candidate predictor measurements and ROC-based measures of proficiency.

b-2-4. Karimi: She suggests recruiting as a possible fix - specifically - select persons who can withstand physical hardships and social pressures of the job.

b-2-5. Davies and Parasuraman: Extroversion seems to be good candidate for vigilance - extroverts are worse. Also, Morris and Gale, 1974 suggest that extroverts generate mental imagery that may interfere with task.

b-2-6. Buckner and McGrath: They proposed that the best predictor of performance in sustained-attention tasks is performance on a sample of the task itself.

b-2-7. SwRI comment on a-10. This could be tested by using existing short tests of NDI methods (9 min. training followed by 15 min. test.

2.3 Many of the tests involve visual inspection, yet there are no specialized eye examinations given to help select/confirm individuals for visual skills.

b-3-1. Michalski et al.: Establish guidelines to require detailed and regular eye exams; German Standard DIN 58220 already does this.

b-3-2. Yonemura: He identifies important parameters in radiography visual tasks and proposes a test based on national norms (test was not yet validated).

b-3-3. SwRI: Such testing would be costly and might not be critical to the Air Force (e.g., some of the disorders discussed in the literature are

age related and the NDI technician is generally relatively young). Therefore, a study should be conducted to determine if this is really a problem before implementing any program.

3.0 Meta-training.

3.1 The general concern that transcends basic and on-the-job training (OJT) is the distribution of training over time. If the goal of NDI training is to produce a proficient "technician," then formal training in theory is too general and lacking in depth. If the goal is an NDI "operator" then the current initial training is probably adequate (if not too intensive) given the time allowed. The goal must be specified before the training can be tailored to accomplish that goal.

c-1-1. SwRI from Chanute visit: Technician candidates should be sent to the field for "vicarious field experience" before attending initial training.

c-1-2. SwRI: Once the role and responsibility questions are answered, training should be adjusted accordingly.

4.0 Initial Training

4.1 The selection/qualification system for NDI instructors needs to be specified/revised. For example, some instructors have no field experience.

d-1-1. SwRI from Chanute visit: Require a minimum of 8 years field experience for instructors and limit their tenure to four years to avoid burn-out.

d-1-2. Lewis et al.: Reduce variation in quality of training - they recommend a Training Review Committee be established to evaluate the Chanute training facility: look at materials, personnel, delivery, equipment, facilities and training aids and make recommendations to ATC.

4.2 Depending on the goal of NDI training, the amount and mix of theory/hands-on experience in the basic course need to be adjusted.

d-2-1. SwRI: Computerized trainer/testing system could assist theory-to-practice and basic-to-OJT integration.

d-2-2. Lewis et al.: Use a two-phase training system aimed at multi-skilled NDI capability with 45/55 class/hands-on ratio in first phase and 35/65 class/hands-on ratio in second phase and both conducted at a centralized facility.

d-2.3. SwRI: The Navy includes "situational awareness" training as part of their NDI basic course to teach students more about how to handle real field situations.

d-2-4. Ainsworth: Should improve training with NDE in mind.

d-2-5. Triggs et al.: In order to get flexible mental representation, a broad diversity of instances must be presented (Posner and Keele, 1968) and proceed from theory to simple practical to difficult practical.

d-2-6. Spanner: Training should include more motor response training.

d-2-7. Embrey: Should train to maximize d', maximize criterion, and ability to adjust criterion to new situation.

d-2-8. Weiner: Use computer-based tutoring devices to provide cuing and knowledge of results.

d-2-9. Triggs et al.: In connection with "pattern training" concept, they describe NDT radiography course by Brock, Wells, and Abrams (1974), in which time went from 80 hours to 11 - started with video information, used "hands-on" for radiographs with detailed knowledge-of-results (KR), tested to determine advancement, and progressed from simple-to-difficult judgments.

4.3 The lack of attrition within the basic course is in sharp contrast to the relatively high failure rate found in many commercial training programs.

d-3-1. Depending on operator/technician decision, this fact might or might not be a concern. If an operator mode is selected, one might expect little attrition. If a technician mode is selected, one would expect the difficulty of the basic course to increase substantially, and (depending on training effectiveness), attrition would also be expected to increase.

4.4 The group-paced mode offers little freedom for advanced or motivated individuals to elaborate their understanding of the information.

d-4-1. Eagle: Develop a computer-based trainer to go into more depth when appropriate.

d-4-2. SwRI: Automated initial training modules would standardize training and free-up instructor time for individual tutoring for both the slow learner and the accelerated learner.

4.5 In a highly "visuospatial" field, there are relatively few visual aids to facilitate understanding (videotapes, films, computer animation, etc.).

d-5-1. SwRI: Develop or purchase or utilize more video-based training materials, especially those that present concrete visuospatial depictions of theoretical constructs.

d-5-2. Eagle: They argue that three-dimensional representation of otherwise invisible processes are a valuable training aid.

d-5-3. Eagle: They argue that the training requirements for UT lend themselves to computer-based training (CBT) and that CBT is the most appropriate training medium.

d-5-4. Triggs et al.: In connection with "pattern training" concept, they describe NDT radiography course by Brock, Wells, and Abrams

(1974), in which time went from 80 hours to 11 - start with video information, hands-on for radiographs with detailed KR, testing to determine advancement and simple-to-difficult judgments.

d-5-5. Spanner et al.: Training should develop mental schemata for broad range of experiences rather than a set of rules.

d-5-6. Triggs et al.: Train internal templates (Annett, 1966) or "perceptual schemata" (Wickens, 1984) or mental models (Gentner and Stevens, 1983).

d-5-7. Triggs et al.: In order to get flexible mental representation, a broad diversity of instances must be presented (Posner and Keele, 1968) and proceed from theory to simple practical to difficult practical.

d-5-8. Bipes et al.: Use self-paced videotape training with real equipment to allow vicarious hands-on training and free instructor's time for those who need it most.

4.6 There is little feedback to initial training instructors from the field.

d-6-1. SwRI: Increase communication and interchange between field and Chanute. Delay the evaluation of initial training until the technician has been in the field at least 6 months. Also, as stated above, require significant field experience for instructors and encourage them to maintain contact with their colleagues in the field.

4.7 The hours of training per technique should be re-assessed. For example, the spectrometric oil analysis program (SOAP) technique (which is highly automated and sometimes not considered to be a NDE procedure) receives 61 hours while ultrasound (a highly complex and theoretical NDE procedure) only receives 51 hours. C-14

d-7-1. Lewis et al.: Reduce variation in quality of training - they recommend a Training Review

Committee be established to evaluate the Chanute training facility: look at materials, personnel, delivery, equipment, facilities and training aids and make recommendations to ATC.

d-7-2. SwRI: Select among operator/technician and generalist/specialist options and then modify initial training accordingly.

d-7-3. SwRI: Because of the highly automated nature of SOAP, it should receive far less time. Because of the highly theoretical nature of UT, eddy current, and radiography, they should receive much more attention (especially if the technician mode is targeted).

4.8 Research is lacking and findings in the current literature have not been assessed to determine what learning components are most important (e.g., theory, motor-behavior, procedural training, pattern-recognition, signal detection in noise, etc.) and which training techniques are most successful (e.g., cuing, knowledge of results, immediate feedback, etc.). Findings are not used to actively update the content/delivery of information in the basic course.

d-8-1. SwRI: Sponsor more applied research in this field. Coordinate with the Air Force Human Resources Laboratory to exploit more about what is known about vigilance and training technology in this area.

d-8-2. Triggs et al.: There is a literature in medical training that suggests that a more "perceptual" approach to training (i.e., specific patterns to look for) can be more useful than theoretical training for the purpose of training technicians in a short period of time (e.g., Alcorn & O'Donnel, 1969).

d-8-3. Triggs et al.: To address motor-training, systematize or train these "movement paths" (Mitchell & Meister, 1973); or provide a jig (Chin Quan, 1974) or more reference points (Triggs et al.) or mechanical guides (Chin Quan & Scott, 1977).

d-8-4. Wheeler et al.: Conduct research and development to determine what training procedures work.

d-8-5. Rummel: Train people on geometry with no flaws.

d-8-6. Williges and North: Knowledge of results or feedback improves performance.

d-8-7. Spanner et al.: Annett (1961) suggests the use of "cuing" to provide immediate or prior hints about what to look for to increase performance.

4.9 There is little evidence supporting the notion that training improves NDI performance.

d-9-1. SwRI: This counter-intuitive finding could be due to the lack of appropriate measures of NDI performance and the manner in which tests have been conducted (i.e., tests of general technicians were conducted on specialized operators). Re-analysis of previous data (where possible) or more appropriate tests will provide a better index of training effectiveness.

4.10 There is too great a discrepancy between initial training and subsequent field environment.

d-10-1. SwRI: Computerized trainer/testing system could assist theory-to-practice integration.

d-10-2. SwRI: The Navy includes "situational awareness" training as part of their NDI basic course to teach students more about how to handle real field situations.

d-10-3. Spanner: Display information should have significant variability - comparable to real world variability.

d-10-4. Karimi: For the non-military NDE technician, she suggests reducing training/field discrepancy by:

a) changing and standardizing management practices in field
b) changing training to be more like field.

d-10-5. SwRI comment from Karimi: Increase continuity through the use of videos in the field to reinforce ethics, standards, etc., and videos in training to provide more realistic expectations.

5.0 On-the-Job Training (OJT)

5.1 There is a lack of standardization, structure, and training materials in the current OJT.

e-1-1. Lewis et al.: Implement rewards and information feedback system.

e-1-2. Summers: Develop and implement computer-based trainer.

e-1-3. SwRI: Develop and implement computer-based training/testing system with psychomotor/theory integrated training and signal detection theory (TSD) measurement built in. Coordinate with ATC and HRL advanced training system to insure that NDI considerations are made.

e-1-4. Eagle: Use computer-based training. Computer-based training (CBT) would allow computer managed instruction to monitor progress and problem areas.

e-1-5. Lewis et al.: They recommend a standardized OJT program be developed and implemented AF-wide including standardized training kits/manuals, test pieces with representative flaws for ultrasound, eddy current surface probe, eddy current bolt hole probe, penetrant, X-ray, and magnetic particle methods.

5.2 The selection, training, qualifications and motivation of the OJT instructor along with his/her organizational relationship with the student need to be standardized/re-assessed.

e-2-1. SwRI Laughlin visit: Specify dedicated instructor responsible for OJT and carefully select and train those instructors.

e-2-2. SwRI Laughlin visit: Use dedicated instructor approach or have different people manage and instruct the technician.

e-2-3. SwRI Laughlin visit: Install "NDI ANSWERS" hotline which any technician can call to get answers to technical questions - no questions asked.

e-2-4. SwRI Laughlin visit: Use dedicated instructor approach or provide other standardized OJT formats.

5.3 There is a lack of diversity in OJT in some instances due to the fact that some installations specialize in certain NDI techniques because of the aircraft serviced.

e-3-1. SwRI: Develop and implement computer-based training/testing system with psychomotor/theory integrated training and TSD measurement built in. Coordinate with ATC and HRL advanced training system to insure that NDI considerations are made.

e-3-2. Eagle: Use computer-based training. Computer-based training (CBT) would allow computer managed instruction to monitor progress and problem areas.

5.4 There is a lack of formal training from the completion of initial training until the optional 7-level course (and that is refresher training).

e-4-1. SwRI: Provide formal and standardized OJT along with tests which are conducted by outside office or by a computerized testing system which reports results directly to ATC or to the NDI Program Office.

5.5 The 7-level course is optional and does not introduce added depth of information.

e-5-1. SwRI: Regardless of operator/technician decision, provide more structured training before and at the -7 level. If the operator mode is selected, then the -7 level course should be very advanced, so that each installation has one technician available for

conducting training, conducting complex/novel inspections, and making final decisions on disputed calls.

e-5-2. Eagle: Save expenses by changing -7 level training to update, refresh, and enhance skills by "using CBT in combination with testing by a single traveling instructor."

5.6 There is no independent critical assessment of NDI proficiency (hands-on or theoretical knowledge) at critical milestones throughout the NDI career progression.

e-6-1. Lewis et al.: There should be an opportunity for periodic re-examination and practice on hardware.

e-6-2. Eagle: A traveling instructor could be used to train NDI (that individual could also systematically conduct impartial tests at certain milestones).

e-6-3. SwRI: Develop and implement computer-based training/testing system with psychomotor/theory integrated training and TSD measurement built in. Coordinate with ATC and HRL advanced training system to insure that NDI considerations are made. Test results can be automatically tele-communicated to ATC or to the NDI Program Office to monitor quality control.

5.7 There is occasionally a lack of update training material and information on new testing equipment delivered to the field.

e-7-1. SwRI: Develop and implement computer-based training/testing system with psychomotor/theory integrated training and TSD measurement built in. Coordinate with ATC and HRL advanced training system to insure that NDI considerations are made.

e-7-2. SwRI: Implement MANPRINT-like procurement requirements to insure that manpower/training/personnel issues are considered in the procurement process.

5.8 There is no vehicle to insure that personnel get adequate practice using infrequent techniques.

e-8-1. Lewis et al.: There should be an opportunity for periodic re-examination and practice on hardware.

e-8-2. Summers: Implement a computer-based training system.

e-8-3. Eagle: Implement a computer-based training system.

e-8-4. SwRI: Develop and implement computer-based training/testing system with psychomotor/theory integrated training and TSD measurement built in. Coordinate with ATC and HRL advanced training system to insure that NDI considerations are made.

6.0 Work Environment

6.1 NDI technicians and their supervisors should be made aware of the effects of environmental variables like temperature, noise, amount of light, humidity, extended time-on-task, and work schedule on human performance.

f-1-1. Herr: Insure that testing occurs in most favorable environment possible. He reports that significant improvement occurred when inspection was moved to quiet room with more light.

f-1-2. Spanner et al.: Utilize rest breaks, feedback, and incentives in NDI. They report that performance is improved if regular rest pauses are provided, knowledge of results (KR) is provided in operational setting, and technician is meaningfully and directly rewarded for quality of work.

f-1-3. Spanner: Reduce time exposure in hot/humid/dangerous environments.

f-1-4. Triggs et al.: Use what we know about variables that influence vigilance: task duration, task complexity, temporal

distribution of signals, signal type, type of feedback, sleep loss, environmental stressors, work-rest cycles, and personality variables (Davies and Tune, 1970; Stroh, 1971).

6.2 There is a lack of research on the effects of work and environmental variables on NDI performance.

f-2-1. Triggs et al.: Conduct more research investigating the effects of environmental variables on NDE performance.

f-2-2. SwRI: Sponsor more research on environmental variables that affect NDI technician performance. Incorporate what facts are already known about environmental factors into OJT and supervisor training modules. Establish guidelines based on what is known.

7.0 Performance

7.1 Because of the nature of NDI tasks, there is an intrinsic lack of performance feedback. For example, there is obviously no real-world immediate feedback for missed calls, and often delayed or no feedback for hits and false alarms. Implications extend to assessment, training, motivation, accountability, and self-confidence.

(Note: see section 9.1 for related solutions.)

g-1-1. Weiner: Use computer-based tutoring devices to provide cuing and knowledge of results.

g-1-2. Lewis et al.: Implement rewards and information feedback system.

g-1-3. Bush et al.: Conduct testing and demonstrations of proficiency. People conducting test must be able to watch display during testing. (SwRI note: a computerized testing apparatus could monitor performance in near real time, determine what information is being presented, determine what the technician should do, determine what the technician did, and provide appropriate feedback to the technician and his/her supervisor.)

g-1-4. Bush et al.: Performance should be judged on basis of importance of the call (e.g., there should be penalties for missing cracks that are large enough to require repair or calling nonexistent crack for repair).

g-1-5. Stone in Rummel: Introduce occasional deliberate flaw (a part-task trainer could do this).

g-1-6. Triggs et al.: Increase immediate feedback. Drury & Addison (1973) found that quality control inspectors' performance increased 50% after given immediate feedback.

g-1-7. Summers: Use computer-based trainer to provide more performance feedback.

g-1-8. Eagle: Use computer-based training to provide more performance feedback.

g-1-9. SwRI: Create a "flaw log" that goes with weapon system but not seen by technician until after inspection.

g-1-10. SwRI: Maintain individual performance graphs showing number of inspections and flaws found and false alarms along with target corresponding values (similar to quality control procedures found in some companies).

g-1-11. SwRI: Devise a way of inserting flaws into everyday NDI for the purpose of providing feedback, measuring proficiency, and providing guidance for training.

g-1-12. Triggs et al.: Providing more measurement and feedback can also detect incompetent, lazy, unethical, or dishonest technicians.

g-1-13. Triggs et al.: Audit and spot check to pick up missed calls.

7.2 There is a lack of systematic research in the area of NDI performance. Also, what is known from other fields is not often applied or communicated to technicians or managers.

g-2-1. SwRI Laughlin visit: Publish more technical information in the Air Force NDI newsletter and subscribe to other technical journals for work area.

g-2-2. Ainsworth: Based on the visual inspection literature, more than one inspection should be conducted.

g-2-3. Triggs et al.: Learn from what related fields are doing to improve performance, especially medical diagnosis.

g-2-4. Triggs et al.: Use what we know from literature on sustained-attention tasks: increase regularity, duration and discriminability of signals; proportion of samples with signals; regular rest pauses; and knowledge of results in operational situation. Meaningfully reward for quality of work and minimize sleep loss and extreme environmental conditions. Use self-paced rather than forced-paced tasks (Holding, 1983).

g-2-5. Lewis et al.: Trade-offs between frequency of inspection and redundant NDI applications should be studied for their impact on reliability.

g-2-6. Lewis et al.: "Risk analysis which uses NDI reliability data should be examined."

g-2-7. Lewis et al.: More analyses of existing data should be conducted before any new data are collected.

g-2-8. Doctor et al.: Conduct more studies of human factors and make multidimensional changes to get real increase in NDE.

7.3 NDI tasks and procedures are often not designed with human performance in mind.

g-3-1. Lewis et al.: We need to pay attention to the man/machine interface, physical/mental attributes, equipment output level/patterns and model transfer functions to account for test object shape.

g-3-2. Ainsworth: NDI tasks should be adapted to optimize human performance.

g-3-3. Triggs et al.: Alter tasks and environment -- the chances of improving performance are greater through alterations in tasks and environment than training and motivation (Harris and Chaney, 1969).

7.4 Nondestructive testing performance is not adequate.

(Obviously, all suggested solutions in this report are targeted at this concern; the following candidate solutions address specific performance issues.)

g-4-1. SwRI from Lewis et al.: Conduct error analysis to guide training development and performance aids.

g-4-2. Sproat: Use redundant inspections with same technique or one search with each of different techniques to get detection up to assumed level.

g-4-3. Davis: Provide a performance-aid like that written by Davis - "Mathematic formulas and references for NDT: Ultrasonics."

g-4-4. Bush et al.: Conduct human-factors research to lead to reduction in human variability.

g-4-5. Cowfer: Data are presented on faults for nuclear components. (SwRI note: there is an implied notion that a database of likely errors and faults should be developed and that training should address identified areas.)

g-4-6. SwRI from Triggs et al.: Create a database of likely faults for training/performance/ engineering-design purposes.

g-4-7. Triggs et al.: Information about where faults are likely increases performance (Chin Quan and Scott, 1977; Forsten and Aaltio, 1982).

g-4-8. Triggs et al.: Improve equipment -- guides should be developed to avoid motor response errors (Chin Quan and Scott, 1977).

g-4-9. Triggs et al.: Alter tasks and environment - the chances of improving performance are greater through alterations in tasks and environment than training and motivation (Harris and Chaney, 1969).

g-4-10. Triggs et al.: Use multiple independent observations to increase accuracy. The mathematics of the expected gain in accuracy have been established (see Green & Swets (1974) and Swets et al., 1979).

g-4-11. Triggs et al.: Adopt and adhere to strict archival records of error, as in radiography.

g-4-12. Sproat: He established a comprehensive data base; this needs to be expanded.

7.5 NDI personnel report barriers to practicing some techniques at some installations.

g-5-1. Forsten and Aaltio: Report increased UT performance with practice; they suggested that 2-3 months practice per year is necessary to keep proficient.

g-5-2. Bush et al.: Conduct testing and demonstrations of proficiency. People conducting test must be able to watch display during testing. (SwRI note: a computerized testing apparatus could monitor performance in near real time, determine what information is being presented, determine what the technician should do, determine what the technician did, and provide appropriate feedback to the technician and his/her supervisor.)

7.6 If the operator approach is taken, then temptations must be resisted to increase demands on the operators.

g-6-1. SwRI: To the extent that the operator role is selected, the responsibility of quality control must be shifted more and more to the NDI Program Office to monitor all functions and assure that TO's maintain extremely high quality and are completely understandable,

equipment is standardized and user-friendly, and demands are not placed on operators which assume technician-level knowledge/skills.

8.0 Procedures

8.1 Most assessments of NDI proficiency have focussed on the detection task. While this is the most important task, there are other human tasks involved. Human error can occur in any tasks and reduce proficiency. No human reliability assessments could be found in which error rates are assessed for all tasks/subtasks. Consequently, even if the existing test results were valid and were perfect indicators of actual field performance, there would still be no true assessment of actual NDI proficiency.

h-1-1. Spanner: Remove data-recording errors by improving training, redesigning/updating forms, standardizing procedures, and using error-checking.

h-1-2. Triggs et al.: Re-design defect recording form, introduce error checking procedure or improve means of identifying position coordinates.

h-1-3. Sproat: Use graphics format to report flaws - to avoid errors in interpreting check-sheets or other methods.

h-1-4. SwRI: Conduct human reliability analyses on various NDI tasks to identify highest sources of errors. Use results of such analyses to guide training, TO development, and equipment purchases.

8.2 To the extent that the Air Force embraces the operator as opposed to the technician approach to NDI, there are implications for the procedures used. Extensive quality control procedures for technical orders must be strictly followed and tested to insure that there is no possibility of misinterpretation. In addition, steps should be taken to minimize possible negative side-effects such as boredom.

h-2-1. Herr: It should be mandatory to identify critical parts and have clear, specific, and detailed procedures for inspecting those parts.

Procedures should be proofed by inspectors under production conditions, with no deviations allowed, and by personnel certified for that part.

h-2-2. Bush et al.: Continue to develop new and better techniques.

h-2-3. Spanner et al.: Improved procedures should be developed and provided as job-performance aids.

h-2-4. Spanner et al.: Use stricter procedures (Reinhardt, 1977; Forsten and Aaltio, 1982) and consider job performance aids.

h-2-5. Triggs et al.: Improve procedures - improve adherence to existing procedures to avoid motor-response errors (Forsten and Aaltio, 1982).

h-2-6. Triggs et al.: Improve procedures - increased detail in procedures leads to improvement in performance (Aaltio and Kauppinen, 1982).

h-2-7. Doctor et al.: Increases in performance due to improved procedures are enhanced by proper training.

h-2-8. SwRI: If the operator mode is selected, it is absolutely critical that TO's be high in quality, understandable, and anticipate most possible variations that could reasonably be expected in the field. During TO development, actual technicians should be used to test and de-bug TO's. The existing system for updating TO's from the field appears to be excellent but could become more formalized by the formation of a database that could archive and help analyze all comments from the field and all reported human errors for each TO.

9.0 Motivation

9.1 The lack of feedback that is characteristic of NDI tasks affects the motivation of the NDI personnel.

(Note: see section 7.1 for related solutions.)

i-1-1. SwRI: Create a "flaw log" that goes with weapon system but not seen by technician until after inspection.

i-1-2. SwRI: Maintain individual performance graphs showing number of inspections and flaws found and false alarms along with target corresponding values (similar to quality control procedures found in some companies).

i-1-3. SwRI from Laughlin visit: Display laboratory graph or sign showing how many "lives saved" or how much money saved due to cost avoidance because of NDI flaws detected.

i-1-4. Lewis et al.: Institute mandatory ratings by supervisors and management.

i-1-5. SwRI: Devise a way of inserting flaws into everyday NDI for the purpose of providing feedback, measuring proficiency, and providing guidance for training.

i-1-6. Summers: Provide a computer-based trainer that gives feedback.

9.2 The effects of response bias and the variables that influence response bias have been ignored in most research on NDI performance proficiency.

i-2-1. Spanner et al.: Help establish appropriate criterion level - overcome pressure of peers and management to alter response bias or decrease quality of inspection by upgrading ethics in the field (Winslow, 1984; McMaster, 1984).

i-2-2. Burley: When management gets right attitude, reliability goes up.

9.3 There is a reported lack of confidence and self-efficacy among nondestructive testing personnel.

i-3-1. Karimi: Decision training/practice could be incorporated in training (SwRI: or provided with a computerized training device).

i-3-2. SwRI: Once an adequate test/measurement system is implemented and corrective training provided contingent upon performance outcomes, self-confidence should dramatically increase.

9.4 The professionalism, feelings of importance in the maintenance community, and general esprit de corps could be improved for NDI.

i-4-1. SwRI from Laughlin visit: Display laboratory graph or sign showing how many "lives saved" or how much money saved due to cost avoidance because of NDI flaws detected.

i-4-2. Lewis et al.: Seek AF approval for and use decals and uniform patches for enhanced esprit.

i-4-3. SwRI: Create and conduct education/public-relations program for educating related technical fields as to what NDI is about and its importance.

i-4-4. Spanner et al.: Provide appropriate criterion level - overcome pressure of peers and management to alter response bias or decrease quality of inspection by upgrading ethics in the field (Winslow, 1984; McMaster, 1984).

9.5 There are reports of low motivation among technicians in the nondestructive testing community.

i-5-1. SwRI from Laughlin visit: Display laboratory graph or sign showing how many "lives saved" or how much money saved due to cost avoidance because of NDI flaws detected.

i-5-2. SwRI: Encourage constructive competition among inspectors - conduct tournaments or publish tallies.

i-5-3. Lewis et al.: Institute mandatory ratings by supervisors and management.

i-5-4. Ward (from manufacturing, not NDI): You can improve quality in manufacturing through a) restructuring, b) increasing accountability - personal graphs of defects produced, and c) competition.

i-5-5. Triggs et al.: To deal with boredom use "job enrichment" (Davis and Cherns, 1975) or use incentive scheme, job rotation, and rest pauses (Davies, Shackleton and Parasuraman, 1983).

i-5-6. SwRI: Once an adequate test/measurement system is implemented and corrective training provided contingent upon performance outcomes, motivation should increase.

10.0 System

10.1 There is a heterogeneity among personnel and NDI tasks across different units, which makes it more difficult to prescribe general, standardized remedies.

j-1-1. Lewis et al.: They found good performance at one installation that: a) was exclusively dedicated to NDI, b) had skilled staff acquired by a selective process, c) performed certification/re-certification at periodic intervals, and d) conducted NDI operations as a specialty.

j-1-2. Lewis et al.: Practical NDI exams should be periodically administered and technicians' performance ratings should be made available to them.

j-1-3. Lewis et al.: Practical exams administered with actual flaws in hardware should be developed for proficiency determination.

j-1-4. Lewis et al.: They recommend that NDI be centralized, and made a full-time certified activity.

j-1-5. Lewis et al.: A formal certification program should be required and consist of both written and practical segments.

j-1-6. Lewis et al.: They recommend a standardized OJT program be developed and implemented AF-wide including standardized training kits/manuals, test pieces with representative flaws for ultrasound, eddy current surface probe, eddy current bolt hole probe, penetrant, X-ray, and magnetic particle methods.

j-1-7. Lewis et al.: NDI should organizationally report to chief of maintenance and personnel.

j-1-8. Lewis et al.: "Depots should be given the authority to monitor field operations, provide well defined NDI procedures as Technical Orders, assure the availability of proper equipment for field use and provide training assistance to the field."

j-1-9. Lewis et al.: Technical order NDI procedures for critical parts should be developed for any new aircraft or major assembly.

j-1-10. Petru: Kelley ALC has taken initiatives to formalize process control procedures and introduce a process assurance program.

j-1-11. Hellier (1984): Exams should include written, specific procedure questions, and practical components (using a library of samples and including preparation and ability to monitor personnel). Tests should be graded by centralized agency.

j-1-12. SwRI: Specification of role and responsibility of NDI personnel should help standardize tasks and expectations. Also, implementation of standardized (computer-based or other) OJT, practice opportunities, and performance tests will facilitate standardization.

10.2 There needs to be more opportunity for practice, especially for infrequently used procedures.

j-2-1. SwRI from Laughlin visit: The system should insure that individuals get more practical exams, even in initial training.

j-2-2. SwRI from Laughlin visit: The system should insure that each individual receives hands-on experience in all methods. Currently, it can be limited in some commands due to the type of tests conducted.

j-2-3. SwRI: A computer-based training, practice, and testing system would provide vicarious practice on infrequent tasks.

10.3 Because of the lack of adequate measures of performance, standards and qualifications for career advancement are less defendable.

j-3-1. Lewis et al.: Redefine standards: "Standards of performance need to be established with the norms set to attainable goals defined from experience."

j-3-2. Triggs et al.: Providing more measurement and feedback can also detect incompetent, lazy, unethical, or dishonest technicians.

j-3-3. SwRI: Adoption of appropriate measures and tests will provide the solution to this concern, making personnel promotion and other decisions much easier and defendable for supervisors.

11.0 Measurement

11.1 The measures of proficiency used in the past by researchers investigating NDI technician performance are not good measures of true detection ability (sensitivity).

k-1-1. Swets: Use relative operator characteristic (ROC) from signal detection theory (TSD) to measure technician proficiency (sensitivity).

k-1-2. Triggs et al.: The common practice of specifying detection probability at a certain level of confidence as a function of defect size needs to be accompanied by the associated false call rates in order to provide a complete specification. Use signal detection measures to assess proficiency. They cite a study that indicates training and continuing experience is related to performance with physicians when using signal detection measures of performance. This contradicts studies using other measures which show no relationship between training and NDI performance.

k-1-3. SwRI: Incorporate signal detection measures of sensitivity into assessment procedures. The general logic of sensitivity and measures from signal detection theory should be taught to all technicians, and incorporated into any NDI automated trainers.

k-1-4. Sproat: Use graphics format to report flaws on tests - to avoid errors in interpreting check-sheets or other methods.

k-1-5. Sproat: Use coefficient of contingency (a normalized chi square approach that compares performance against chance).

k-1-6. Boisvert et al.: They cite Packman and Malpani (1978) who proposed subtractive penalty on false calls, and Sproat and Sharpe (1979) who suggest coefficient of contingency for measuring proficiency.

k-1-7. Wheeler et al.: Use ROC from TSD as proficiency measure because it works very well for describing and analyzing technician proficiency.

k-1-8. Glasch: Instead of older POD measures, use ROC.

k-1-9. Triggs et al.: Investigate improved TSD procedures - Hanley and McNeil (1983) have developed an approach to increase statistical sensitivity of comparing ROC's.

k-1-10. Triggs et al.: Use the TSD Rating-Scale Method with 5 levels, and discount contingency coefficient and other measures.

k-1-11. Triggs et al.: Use specialized methods (e.g., Watson and Nichols, 1976) for applying signal detection theory to vigilance tasks like UT (i.e., where individual observation episodes are not defined).

k-1-12. Spanner et al.: Should use ROC from TSD to assess technician performance.

11.2 The measures of proficiency used in the past by researchers investigating NDI technician performance are not good measures of the technician's response bias.

k-2-1. Swets: Use the relative operating characteristic (ROC) from signal detection theory (TSD) to measure effects on technician's performance due to response bias (location of response criterion).

k-2-2. SwRI: Incorporate signal detection measures of response bias into assessment procedures. The general logic of response bias should be taught to all technicians, and incorporated into any NDI automated trainers. Part of each TO should specify the response bias appropriate for a given test piece.

11.3 The experimental design or approach taken in the past by researchers investigating NDI technician performance were sometimes inappropriate or led to misinterpretation.

k-3-1. Boisvert et al.: The appropriate practical exam would include the following features: samples must reflect typical geometry; flaws must produce realistic responses; flaw size range must be large and representative; number of test opportunities must be large enough for statistical validity; and ratio of un-flawed to flawed must be realistic (however economics dictate compromise).

k-3-2. SwRI: A generic test protocol should be developed and implemented. Depending on the outcome of the role/responsibility decisions, very different protocols could result. The protocol should take TSD measurement into account and satisfy all criticisms of previous test protocols in the literature.

11.4 Because of inappropriate measures of proficiency, there might be inaccuracies in the literature with regard to potential correlates of NDI proficiency.

k-4-1. SwRI: All conclusions based on earlier inadequate measures of proficiency and how it relates to selection variables and training variables must be questioned. If possible, re-analysis should be conducted. Any new research should take these considerations into account.

11.5 The lack of an appropriate measure of proficiency has potential wide-ranging negative effects ranging from difficulty in training to low motivation/morale.

k-5-1. Boisvert et al.: A defendable proficiency measure would help in many ways including the establishment of training needs, tailored to the proficiency of individual.

k-5-2. Boisvert et al.: A defendable proficiency measure would allow better assignment of personnel to tasks (e.g., the most proficient technicians should be the OJT instructors).

k-5-3. Boisvert et al.: A defendable proficiency measure would allow accurate feedback and motivation to improve.

k-5-4. SwRI from Wheeler et al.: Use self-efficacy in measurement scheme.

k-5-5. SwRI: Implementation of TSD measurement, immediate feedback, and standardized OJT should address this concern.

12.0 Equipment

12.1 Equipment should be selected which is semi-automated and reduces the number and complexity of the decisions made by the inspector (especially if the NDI operator path is taken).

1-1-1. Ainsworth: Automate human out of system.

1-1-2. Burley: Whenever possible, automated NDE should replace manual NDE.

1-1-3. Lewis et al.: Automatize the procedure and the control functions susceptible to error.

1-1-4. Lewis et al.: Improve equipment through automation and digital processors.

1-1-5. Lewis et al.: Combine search and flaw characterization functions in the same equipment.

1-1-6. Bush et al.: Increase automation - automated UT equipment offers several advantages, one is that raw data can be taken and, thus, the possibility of archiving data and using new interpretations as they are invented. Also, features can be correlated with previously associated types, sizes, shapes, and orientations of defects. (Individual records of geometry could be taken and, subsequently, software could look for changes.)

1-1-7. Burley: There needs to be more quantification built into the equipment, especially to assist in sizing.

1-1-8. SwRI: Automation suits the operator mode more than the technician mode. Consideration must be made of the maintenance implications of some highly automated equipment.

12.2 NDI equipment (displays, controls, and sensors) should be standardized as much as possible (especially if the NDI operator path is taken).

1-2-1. Lewis et al.: Permanent record and visual displays are necessary adjuncts.

1-2-2. Wheeler et al.: Conduct further analyses of equipment design and establish standards for equipment.

1-2-3. Triggs et al.: Capitalize on methods from other disciplines; for example, Berger (1976)-draw solutions from clinical chemistry (e.g., standardize calibration procedures and adopt standard reference materials).

1-2-4. SwRI: Standardization is more important in the operator mode than in the technician mode.

12.3 NDI equipment displays, controls, and sensors should be designed, evaluated, and selected on the basis of their ease of use among actual NDI technicians.

1-3-1. Lewis et al.: The signal-to-noise ratio should be increased in detection and readout functions.

1-3-2. Lewis et al.: Promotes operator vigilance through enhanced stimulation to sensual perception.

1-3-3. Lewis et al.: Channel the operator's attention to flaw indications.

1-3-4. Ainsworth: Design equipment with human in mind.

1-3-5. Spanner et al.: Human-engineering guidelines should be developed and given to equipment manufacturers.

1-3-6. Wheeler et al.: Manufacturers should: simplify controls to only essentials; increase size of controls for easier use with gloves; increase size of display for better viewing from distance; improve anti-glare for display; design search units with non-slip surface for gloves; insure that provisions are made to connect two units together in master/slave arrangement; incorporate voice communication in equipment; and improve cabling for master/slave.

1-3-7. Herr: States that direct reading digital equipment and portable equipment reduce fatigue and help the NDI technician.

12.4 Because of the criticality of the technical orders in Air Force NDI, great care must be taken to insure that all technical orders correspond to current equipment available.

1-4-1. SwRI: Standardization, automation, and user friendliness are critical equipment goals if the operator mode is selected, less critical if the technician mode is selected.

12.5 There is a lack of knowledge and research in the areas of display design, controls design, and sensor design to maximize human performance.

1-5-1. Triggs et al.: Conduct research to design optimal probe, feedback about pressure, etc.

12.6 Weapon-system manufacturers do not always consider NDI testing when designing a new system; NDI equipment manufacturers are not always accessible.

1-6-1. Lewis et al.: Provide positive assurance that the equipment is working properly.

1-6-2. Spanner et al.: Establish qualifying systems.

1-6-3. Spanner et al.: All equipment should be qualified by demonstration.

1-6-4. Herr: Institute better quality control-Convair has developed some standard receiving acceptance tests and are finding a high rejection rate. Standardization is being improved by Air Force Materials Lab, industry *associations*, technical societies, and National Bureau of Standards.

1-6-5. Spanner et al.: Human-engineering guidelines should be developed and given to equipment manufacturers.

1-6-6. Wheeler et al.: Manufacturers should: simplify controls to only essentials; increase size of controls for easier use with gloves; increase size of display for better viewing from distance; improve anti-glare for display; design search units with non-slip surface for gloves; insure that provisions are made to connect two units together in master/slave arrangement; incorporate voice communication in equipment; and improve cabling for master/slave.

1-6-7. Herr: Design original equipment for ultrasonic inspection (e.g., use only flat surfaces, square corners, etc.).

1-6-8. Burley: We need to introduce NDE into design stages of equipment.

V. RESULTS OF PHASE III: RECOMMENDATIONS

In Phase III, identified solutions for each area of concern were condensed into one to three recommendations. To provide an estimate of the promise, feasibility, and cost of each recommendation, an expert panel was convened. A modified Delphi approach was utilized (as described in the METHODS section).

The recommendations and the statistics resulting from the paneling exercise are presented on the following pages. Some of the recommendations (especially those dealing with automated training/measurement devices) might appear to be redundant, however they are not. In some cases, a specific feature of such a device might be highlighted to address a specific area of concern. In other cases, although the wordings of two recommendations might be identical, their merit was being evaluated for addressing different areas of concern.

At the top of each page, the relevant area of concern is stated to provide the reader the appropriate context. Next, the recommendation is stated. The recommendations presented here include any minor wording changes and clarifications that emerged during the paneling exercise. If more than one recommendation was considered for a specific concern, they appear on subsequent pages.

Below each recommendation, statistics are presented which describe the outcome of the panel's evaluation. Results are presented for promise, feasibility, cost, promise plus feasibility, and promise plus feasibility plus cost (respectively). Below each of the three primary scales, a frequency distribution of ratings is presented. In all cases, the distribution presented represents the last rating taken for that recommendation. An asterisk to the right of the distribution signifies that there was disagreement (range greater than three) among the panelists on the first pass. The reader can assess whether the disagreement was resolved in the second pass by evaluating the final distribution. In most cases, disagreements were resolved. The reader is reminded that only one pass was made for the cost estimations.

In addition to the frequency distribution, mean ratings for the panel are also presented. These means were weighted by the confidence ratings reported by each panelist for that rating. The weighted mean provides the reader with an estimate of how the panel perceived the recommendation with regard to promise, feasibility, and cost.

In addition to the weighted means, standard scores are presented to assist the reader evaluate how a specific recommendation was rated relative to all other recommendations considered in this exercise. The standard score yields an overall mean of 50 and a standard deviation of 10 across all recommendations. Consequently, any recommendations with standard scores greater than 50 were rated higher than average for the trait being considered. For example, if the standard score for "feasibility" of a specific recommendation was equal to 65, then that recommendation was 1.5 standard deviations above the mean feasibility for all recommendations. Standard scores would rarely be expected to be less than 30 or greater than 70.

The two composite scores presented at the bottom of each page provide overall estimates of the merit of a given recommendation. Although each should be considered, it is recommended that the promise plus feasibility be given greater consideration, primarily because of the many variables that could potentially influence cost or the panelists' perception of cost.

1.0 Global Areas of Concern

1.1 The Measurement of NDI proficiency.

A. Adopt Signal Detection Theory (TSD) methodology as a standard measurement system for NDI technician proficiency. This involves educating the NDI community about the general principles of TSD and the definitions and implications of the two measures (sensitivity and response bias).

Solution Shows Promise?

Not Promising				Very Promising
1	2	3	4	5
			5	1

Weighted Mean- 4.24
Standard Score- 52.4

Solution is Attainable?

Not Feasible				Very Feasible
1	2	3	4	5
		2	3	1

Weighted Mean- 3.95
Standard Score- 52.8

Estimated Cost of Solution?

Very High				Very Low
1	2	3	4	5
	1	1	4	

Weighted Mean- 3.75
Standard Score- 57.3

Promise + Feasibility - 8.19
Standard Score - 53.2

Promise + Feasibility + Cost - 11.94
Standard Score - 56.9

1.2 The <u>Role</u> of Air Force NDI personnel: operator or technician?

and

1.3 The <u>Responsibility</u> of Air Force NDI personnel: general responsibility for all NDI techniques or specialist in certain techniques.

A. Compromise between the general/specific and operator/technician extremes. Take steps to ensure that 3-level "operators" can proceed to competent 7-level "technicians."

<u>Solution Shows Promise</u>?

Not Promising 1	2	3	4	Very Promising 5
		1	1	4

Weighted Mean= 4.52
Standard Score= 59.0

<u>Solution is Attainable</u>?

Not Feasible 1	2	3	4	Very Feasible 5
		2	4	*

Weighted Mean= 3.73
Standard Score= 49.0

<u>Estimated Cost of Solution</u>?

Very High 1	2	3	4	Very Low 5
	4	1	1	

Weighted Mean= 2.50
Standard Score= 40.1

Promise + Feasibility = 8.25
Standard Score = 53.9

Promise + Feasibility + Cost = 10.75
Standard Score = 46.5

Specific Areas of Concern and Possible Solutions

2.0 Selection

2.1 In a highly technical area, there is no intentional selection mechanism. NDI technicians come from the general manpower pool. Quality of performance could probably be enhanced by selecting from the electronics or mechanics personnel pools.

A. Select better NDI candidates. As a short-term partial solution, select samples from the population who are higher in electronic and mechanical abilities, measure and compare their success with personnel selected using the current approach. (Coordinate this effort with the Air Training Command, Human Resources Laboratory, and Occupational Measurement Center at Randolph AFB.)

Solution Shows Promise?

Not Promising 1	2	3	4	Very Promising 5
			1	5

Weighted Mean= 4.89
Standard Score= 67.9

Solution is Attainable?

Not Feasible 1	2	3	4	Very Feasible 5
		1	2	3

Weighted Mean= 4.50
Standard Score= 62.4

Estimated Cost of Solution?

Very High 1	2	3	4	Very Low 5
	1	1	1	3 *

Weighted Mean= 4.19
Standard Score= 63.3

Promise + Feasibility = 9.39
Standard Score = 67.9

Promise + Feasibility + Cost = 13.58
Standard Score = 71.1

B. Do not select NDI personnel from initial recruits. Rather, select from persons with Air Force electronics or Fabrications Branch experience (as the Navy does), so that performance history and experience can be used as a predictor and so that personnel go through initial training with an existing knowledge base for Air Force maintenance procedures. Test a sample of individuals using this approach to insure that the selection/training power gained from having related field experience more than compensates for any possible biases such individuals might have picked up in earlier assignments.

Solution Shows Promise?

Not Promising				Very Promising	
1	2	3	4	5	
		1	4	1	*

Weighted Mean- 4.10
Standard Score- 49.0

Solution is Attainable?

Not Feasible				Very Feasible	
1	2	3	4	5	
	3	2	1		*

Weighted Mean- 2.54
Standard Score- 28.1

Estimated Cost of Solution?

Very High				Very Low
1	2	3	4	5
	2	4		

Weighted Mean- 2.60
Standard Score- 41.5

Promise + Feasibility - 6.64
Standard Score - 34.2

Promise + Feasibility + Cost - 9.24
Standard Score - 33.4

2.2 There are a number of candidate selection variables proposed in the technical literature, but little systematic research into which variables predict good NDI technicians. The literature that is available is not meaningful since measures of the predicted variable (proficiency) were not adequate.

A. For a potentially better, long-range solution, coordinate with the Human Resources Laboratory, and conduct research to establish predictors of proficient NDI personnel. Provide them with the suggestions for selection criteria found in the literature and presented in this report. Consider task-performance-related selection (i.e., selection on the basis of performance on actual representative tasks rather than on paper and pencil or other psychometric tests).

Solution Shows Promise?

Not Promising 1	2	3	4	Very Promising 5
		2	3	1

Weighted Mean- 4.22
Standard Score- 51.9

Solution is Attainable?

Not Feasible 1	2	3	4	Very Feasible 5
	1	3	2	

Weighted Mean- 3.11
Standard Score- 38.1

Estimated Cost of Solution?

Very High 1	2	3	4	Very Low 5
	2	3	1	

Weighted Mean- 2.71
Standard Score- 43.0

Promise + Feasibility - 7.33
Standard Score - 42.7

Promise + Feasibility + Cost - 10.04
Standard Score - 40.3

B. To the extent possible, analyze previous proficiency data to provide appropriate judgments about what variables predict proficiency, using sensitivity from TSD as the measure of proficiency.

<u>Solution Shows Promise</u>?

Not Promising 1	2	3	4	Very Promising 5
	2	1	3	

Weighted Mean- 3.42
Standard Score- 32.9

<u>Solution is Attainable</u>?

Not Feasible 1	2	3	4	Very Feasible 5	
			4	2	*

Weighted Mean- 4.35
Standard Score- 59.8

<u>Estimated Cost of Solution</u>?

Very High 1	2	3	4	Very Low 5
		2	3	1

Weighted Mean- 3.79
Standard Score- 57.8

Promise + **Feasibility** - 7.77
Standard Score - 48.0

Promise + **Feasibility** + **Cost** - 11.56
Standard Score - 53.6

2.3 Many of the tests involve visual inspection, yet there are no specialized eye examinations given to help select/confirm individuals for visual skills.

A. Since such testing could be costly, a study should be conducted in coordination with the Clinical Sciences Division at the School of Aerospace Medicine. In this study, a cross section of NDI technicians of different ages should be tested for the various deficiencies identified in the literature. If the study determines that a problem exists, then mandatory eye exams should be required at a schedule commensurate with known developmental incidence rates (more frequent testing as the technician ages).

Solution Shows Promise?

Not Promising				Very Promising	
1	2	3	4	5	
		1	3	2	*

Weighted Mean= 4.19
Standard Score= 51.2

Solution is Attainable?

Not Feasible				Very Feasible
1	2	3	4	5
			3	3

Weighted Mean= 4.59
Standard Score= 64.0

Estimated Cost of Solution?

Very High				Very Low
1	2	3	4	5
		2	4	

Weighted Mean= 3.63
Standard Score= 55.6

Promise + Feasibility = 8.78
Standard Score = 60.4

Promise + Feasibility + Cost = 12.41
Standard Score = 61.0

3.0 Meta-training

3.1 The general concern that transcends basic and on-the-job training (OJT) is the distribution of training over time. If the goal of NDI training is to produce a proficient "technician," then formal training in theory is too general and lacking in depth. If the goal is an NDI "operator," then the current initial training is probably adequate (if not too intensive) given the time allowed. The goal must be specified before the training can be tailored to accomplish that goal.

A. The Air Force has established that an operator is an interim goal and that a technician is the ultimate goal of their NDI Program. Consequently, the general structure of NDI training should be directed at that goal. Specifically, level 3, level 5, and level 7 training all need to be intensified with more content, depth, feedback, testing, etc.

<u>Solution Shows Promise</u>?

Not Promising 1	2	3	4	Very Promising 5
			3	3

Weighted Mean- 4.57
Standard Score- 60.2

<u>Solution is Attainable</u>?

Not Feasible 1	2	3	4	Very Feasible 5
		3	3	*

Weighted Mean- 3.59
Standard Score- 46.5

<u>Estimated Cost of Solution</u>?

Very High 1	2	3	4	Very Low 5
1	5			

Weighted Mean- 1.78
Standard Score- 30.2

Promise + Feasibility - 8.16
Standard Score - 52.8

Promise + Feasibility + Cost - 9.94
Standard Score - 39.5

B. Technician candidates should be sent to the field for "vicarious field experience" or selected from a population who already have some real-world maintenance experience (as in the Navy) before attending initial NDI training.

Solution Shows Promise?

Not Promising 1	2	3	4	Very Promising 5
		1	4	1

Weighted Mean= 4.00
Standard Score= 46.7

Solution is Attainable?

Not Feasible 1	2	3	4	Very Feasible 5
	1	3	2	

Weighted Mean= 3.31
Standard Score= 41.6

Estimated Cost of Solution?

Very High 1	2	3	4	Very Low 5	
	3	1	1	1	*

Weighted Mean= 3.12
Standard Score= 48.6

Promise + Feasibility = 7.31
Standard Score = 42.4

Promise + Feasibility + Cost = 10.43
Standard Score = 43.7

4.0 Initial Training

4.1 The selection/qualification system for NDI instructors needs to be specified/revised. For example, some instructors have no field experience.

A. Increase requirements for instructor selection. Require a minimum of 5-6 years field experience for instructors and limit their tenure to four years to avoid burn-out. Increase train-the-trainer materials. Make every attempt to select instructors who are motivated to teach (e.g., allow personnel to turn down or volunteer instructor assignments).

<u>Solution Shows Promise</u>?

Not Promising				Very Promising
1	2	3	4	5
			1	5

Weighted Mean- 4.85
Standard Score- 66.9

<u>Solution is Attainable</u>?

Not Feasible				Very Feasible
1	2	3	4	5
			2	4

Weighted Mean- 4.80
Standard Score- 67.7

<u>Estimated Cost of Solution</u>?

Very High				Very Low	
1	2	3	4	5	
	3		2	1	*

Weighted Mean- 3.33
Standard Score- 51.5

Promise + Feasibility - 9.65
Standard Score - 71.0

Promise + Feasibility + Cost - 12.98
Standard Score - 65.9

4.2 Depending on the goal of NDI training, the amount and mix of theory/hands-on experience in the initial course need to be adjusted.

A. Develop and use a computerized trainer/testing system to a) provide more accurate and immediate feedback, b) assist theory-to-practice and basic-to-OJT integration, c) provide visual representations of complex physical phenomena, d) provide more hands-on experience during initial training, and e) provide more theoretical training during OJT. Use the trainer to augment, not replace, the instructor. Efforts should be coordinated with the Training Systems Division of the Human Resources Laboratory and the Advanced Training System Branch at the Air Training Command.

Solution Shows Promise?

Not Promising 1	2	3	4	Very Promising 5
			2	4

Weighted Mean= 4.71
Standard Score= 63.6

Solution is Attainable?

Not Feasible 1	2	3	4	Very Feasible 5	
		2	4		*

Weighted Mean= 3.69
Standard Score= 48.2

Estimated Cost of Solution?

Very High 1	2	3	4	Very Low 5
2	2	2		

Weighted Mean= 1.83
Standard Score= 30.9

Promise + Feasibility = 8.40
Standard Score = 55.8

Promise + Feasibility + Cost = 10.23
Standard Score = 42.0

4.3 The lack of attrition within the initial course is in sharp contrast to the relatively high failure rate found in many commercial training programs.

A. Increase the selection process within initial technical training. Allow personnel who do not pass to recycle or move into another specialty code.

Solution Shows Promise?

Not Promising 1	2	3	4	Very Promising 5
		2	2	2

Weighted Mean= 4.29
Standard Score= 53.6

Solution is Attainable?

Not Feasible 1	2	3	4	Very Feasible 5	
		3	3		*

Weighted Mean= 3.50
Standard Score= 44.9

Estimated Cost of Solution?

Very High 1	2	3	4	Very Low 5	
	1	3	1	1	*

Weighted Mean= 3.38
Standard Score= 52.2

Promise + Feasibility = 7.79
Standard Score = 48.3

Promise + Feasibility + Cost = 11.17
Standard Score = 50.2

4.4 The group-paced mode offers little freedom for advanced or motivated individuals to elaborate their understanding of the information.

A. Move toward a pace that best fits the average student, not the slower student. Implement a computer-based trainer or other additional training materials to provide advanced instruction for the gifted trainees and remedial instruction for the trainee having difficulty.

Solution Shows Promise?

Not Promising 1	2	3	4	Very Promising 5
			2	4

Weighted Mean= 4.70
Standard Score= 63.3

Solution is Attainable?

Not Feasible 1	2	3	4	Very Feasible 5
		1	3	2

Weighted Mean= 4.35
Standard Score= 59.8

Estimated Cost of Solution?

Very High 1	2	3	4	Very Low 5
		3	1	2

Weighted Mean= 3.89
Standard Score= 59.2

Promise + Feasibility = 9.05
Standard Score = 63.7

Promise + Feasibility + Cost = 12.94
Standard Score = 65.6

4.5 In a highly "visuospatial" field, there are relatively few visual aids to facilitate understanding (videotapes, films, computer animation, etc.).

A. Develop or purchase or utilize more video-based training materials, especially those that present dynamic concrete visuospatial depictions of theoretical constructs.

<u>Solution Shows Promise</u>?

Not Promising 1	2	3	4	Very Promising 5	
			4	2	*

Weighted Mean- 4.39
Standard Score- 56.0

<u>Solution is Attainable</u>?

Not Feasible 1	2	3	4	Very Feasible 5
			3	3

Weighted Mean- 4.48
Standard Score- 62.1

<u>Estimated Cost of Solution</u>?

Very High 1	2	3	4	Very Low 5	
	1	3	1	1	*

Weighted Mean- 3.38
Standard Score- 52.2

Promise + Feasibility - 8.87
Standard Score - 61.5

Promise + Feasibility + Cost - 12.25
Standard Score - 59.6

4.6 There is little feedback to initial training instructors from the field.

A. Increase communication and interchange among the field laboratories and Chanute. Require instructors to visit the field periodically. Include articles/bulletins from the field and from Chanute in the AF NDI publication. Delay the evaluation of initial training until the technician has been in the field at least 3 months. Also, require significant field experience for instructors and encourage them to maintain contact with their colleagues in the field.

<u>Solution Shows Promise</u>?

Not Promising 1	2	3	4	Very Promising 5
		1		5

Weighted Mean- 4.83
Standard Score- 66.4

<u>Solution is Attainable</u>?

Not Feasible 1	2	3	4	Very Feasible 5
		1	1	4

Weighted Mean- 4.68
Standard Score- 65.6

<u>Estimated Cost of Solution</u>?

Very High 1	2	3	4	Very Low 5
			4	2

Weighted Mean- 4.38
Standard Score- 65.9

Promise + Feasibility - 9.51
Standard Score - 69.3

***Promise* + Feasibility + Cost - 13.89**
Standard Score - 73.8

4.7 The hours of training per technique should be re-assessed. For example, the spectrometric oil analysis program (SOAP) technique (which is highly automated and sometimes not considered to be an NDE procedure) receives 61 hours while ultrasound (a highly complex and theoretical NDE procedure) only receives 51 hours.

A. Re-evaluate distribution of training. More time should be spent on difficult tests. Because a general technician is the goal of training, more depth must be provided at some point. However, persons in initial training are already receiving a great deal of information in a relatively short period of time. Therefore, the 5-Level training, Advanced Course, and other supplemental training along with OJT must be made more formal and intense.

Solution Shows Promise?

Not Promising 1	2	3	4	Very Promising 5
		1	2	3

Weighted Mean- 4.55
Standard Score- 59.8

Solution is Attainable?

Not Feasible 1	2	3	4	Very Feasible 5
		2	3	1

Weighted Mean- 3.95
Standard Score- 52.8

Estimated Cost of Solution?

Very High 1	2	3	4	Very Low 5
		3	2	1

Weighted Mean- 3.60
Standard Score- 55.2

Promise + Feasibility - 8.50
Standard Score - 57.0

Promise + Feasibility + Cost - 12.10
Standard Score - 58.3

4.8 Research is lacking and findings in the current literature have not been assessed to determine what learning components are most important (e.g., theory, motor-behavior, procedural training, pattern-recognition, signal detection in noise, etc.) and which training techniques are most successful (e.g., cuing, knowledge of results, immediate feedback, etc.). Findings are not used to actively update the content/delivery of information in the initial course.

A. The Air Force should sponsor more applied research in this field. Coordinate with the Air Force Human Resources Laboratory to exploit more about what is known about vigilance and training technology in this area. The literature offers many promising topics. Provide such topics (as provided in the Areas of Concern and Possible Solution sections of this report) to participating researchers.

<u>Solution Shows Promise</u>?

Not Promising 1	2	3	4	Very Promising 5	
		2	4		*

Weighted Mean= 3.65
Standard Score= 38.3

<u>Solution is Attainable</u>?

Not Feasible 1	2	3	4	Very Feasible 5
	1	1	4	

Weighted Mean= 3.53
Standard Score= 45.4

<u>Estimated Cost of Solution</u>?

Very High 1	2	3	4	Very Low 5
	3	3		

Weighted Mean= 2.33
Standard Score= 37.8

Promise + Feasibility = 7.18
Standard Score = 40.8

Promise + Feasibility + Cost = 9.51
Standard Score = 35.7

4.9 There is little evidence supporting the notion that training improves NDI performance.

A. Re-analyze previous data (where possible) or conduct more appropriate tests to provide a better index of training effectiveness (coordinate with the Human Resources Laboratory to conduct such tests).

<u>Solution Shows Promise</u>?

Not Promising 1	2	3	4	Very Promising 5	
		3	3		*

Weighted Mean= 3.47
Standard Score= 34.0

<u>Solution is Attainable</u>?

Not Feasible 1	2	3	4	Very Feasible 5	
			6		*

Weighted Mean= 4.00
Standard Score= 53.7

<u>Estimated Cost of Solution</u>?

Very High 1	2	3	4	Very Low 5
		2	2	2

Weighted Mean= 4.14
Standard Score= 62.6

Promise + Feasibility = 7.47
Standard Score = 44.4

Promise + Feasibility + Cost = 11.61
Standard Score = 54.0

4.10 There is too great a discrepancy between initial training and subsequent field environment.

A. Develop and utilize a series of videotapes or interactive courseware designed to provide information about what the field is like.

Solution Shows Promise?

Not Promising 1	2	3	4	Very Promising 5	
		1	4	1	*

Weighted Mean- 4.05
Standard Score- 47.9

Solution is Attainable?

Not Feasible 1	2	3	4	Very Feasible 5
			3	3

Weighted Mean- 4.58
Standard Score- 63.8

Estimated Cost of Solution?

Very High 1	2	3	4	Very Low 5	
	1	2	2	1	*

Weighted Mean- 3.44
Standard Score- 53.0

Promise + Feasibility - 8.63
Standard Score - 58.6

Promise + Feasibility + Cost - 12.07
Standard Score - 58.0

5.0 On-the-Job Training (OJT)

5.1 There is a lack of standardization, structure, and training materials in the current OJT.

A. Develop and implement computer-based training/testing system with psychomotor/theory integrated training and signal detection theory (TSD) measurement built in. Build in tele-communications or locked databases either of which could be used to provide on-site supervisors, ATC headquarters, MAJCOM's, or the NDI Program Office information about the system's use and the technicians' performance. Coordinate with the Training Systems Division of the Human Resources Laboratory and the Advanced Training System Branch at the Air Training Command to ensure that NDI considerations are made for automated trainers being developed.

Solution Shows Promise?

Not Promising 1	2	3	4	Very Promising 5
			3	3

Weighted Mean= 4.61
Standard Score= 61.2

Solution is Attainable?

Not Feasible 1	2	3	4	Very Feasible 5
		1	4	1

Weighted Mean= 4.14
Standard Score= 56.1

Estimated Cost of Solution?

Very High 1	2	3	4	Very Low 5
2	1	3		

Weighted Mean= 2.07
Standard Score= 34.2

Promise + Feasibility = 8.75
Standard Score = 60.0

Promise + Feasibility + Cost = 10.82
Standard Score = 47.1

B. Provide OJT instructors more standardized course material (including weapon-specific information) during OJT. To the extent possible, adjust duty assignments to ensure that OJT instructors have adequate time to devote to training.

<u>Solution Shows Promise</u>?

Not Promising				Very Promising
1	2	3	4	5
		2	**1**	**3**

Weighted Mean- 4.39
Standard Score- 56.0

<u>Solution is Attainable</u>?

Not Feasible				Very Feasible
1	2	3	4	5
	1	**1**	**4**	*

Weighted Mean- 3.63
Standard Score- 47.2

<u>Estimated Cost of Solution</u>?

Very High				Very Low
1	2	3	4	5
		3	**2**	**1**

Weighted Mean- 3.89
Standard Score- 59.2

Promise + Feasibility - 8.02
Standard Score - 51.1

Promise + Feasibility + Cost - 11.91
Standard Score - 56.6

C. If a computerized, tele-communication system is not possible, then establish evaluation teams or individuals who periodically visit NDI labs and evaluate each technician's performance.

<u>Solution Shows Promise</u>?

Not Promising				Very Promising	
1	2	3	4	5	
		3	3		*

Weighted Mean= 3.55
Standard Score= 36.0

<u>Solution is Attainable</u>?

Not Feasible				Very Feasible	
1	2	3	4	5	
	2	3	1		*

Weighted Mean= 2.83
Standard Score= 33.2

<u>Estimated Cost of Solution</u>?

Very High				Very Low
1	2	3	4	5
	4	2		

Weighted Mean= 2.25
Standard Score= 36.7

Promise + Feasibility = 6.38
Standard Score = 31.0

Promise + Feasibility + Cost = 8.63
Standard Score = 28.1

5.2 The selection, training, qualifications, and motivation of the OJT instructor along with his/her organizational relationship with the student need to be standardized/ re-assessed.

A. Specify a dedicated instructor responsible for OJT and carefully select and train that individual (i.e., 7-level technician who wants to be instructor). Make such technicians dedicated instructors or greatly reduce other responsibilities.

<u>Solution Shows Promise</u>?

Not Promising				Very Promising	
1	2	3	4	5	
		1	5		*

Weighted Mean= 3.86
Standard Score= 43.3

<u>Solution is Attainable</u>?

Not Feasible				Very Feasible	
1	2	3	4	5	
	1	4	1		*

Weighted Mean= 2.92
Standard Score= 34.8

<u>Estimated Cost of Solution</u>?

Very High				Very Low
1	2	3	4	5
		1	3	2

Weighted Mean= 4.19
Standard Score= 63.3

Promise + Feasibility = 6.78
Standard Score = 35.9

Promise + Feasibility + Cost = 10.97
Standard Score = 48.4

B. Install "NDI ANSWERS" hotline or electronic bulletin board on which any technician can get information and answers to technical questions - no questions asked.

<u>Solution Shows Promise</u>?

Not Promising				Very Promising	
1	2	3	4	5	
	1	1	3	1	*

Weighted Mean= 3.71
Standard Score= 39.8

<u>Solution is Attainable</u>?

Not Feasible				Very Feasible
1	2	3	4	5
		1	3	2

Weighted Mean= 4.29
Standard Score= 58.8

<u>Estimated Cost of Solution</u>?

Very High				Very Low
1	2	3	4	5
		1	2	3

Weighted Mean= 4.50
Standard Score= 67.6

Promise + Feasibility = 8.00
Standard Score = 50.9

Promise + Feasibility + Cost = 12.50
Standard Score = 61.7

5.3 There is a lack of diversity in OJT in some instances due to the fact that some installations specialize in certain NDI techniques because of the aircraft they service.

A. Develop and implement a computer-based training/testing system with psychomotor/theory integrated training and TSD measurement built in. This system should eventually offer courseware ranging from beginner to advanced levels for all techniques expected of AF NDI personnel. Use this system to periodically test NDI personnel on all techniques to determine proficiency/weakness. Coordinate with the Training Systems Division of the Human Resources Laboratory and the Advanced Training System Branch at the Air Training Command to ensure that NDI considerations are made for automated trainers being developed.

<u>Solution Shows Promise</u>?

Not Promising				Very Promising
1	2	3	4	5
			2	4

Weighted Mean- 4.72
Standard Score- 63.8

<u>Solution is Attainable</u>?

Not Feasible				Very Feasible
1	2	3	4	5
		1	5	

Weighted Mean- 3.85
Standard Score- 51.0

<u>Estimated Cost of Solution</u>?

Very High				Very Low
1	2	3	4	5
2	2	2		

Weighted Mean- 1.82
Standard Score- 30.8

Promise + Feasibility - 8.57
Standard Score - 57.8

Promise + Feasibility + Cost - 10.39
Standard Score - 43.4

5.4 There is a lack of formal training from the completion of initial training until the optional Advanced Course, and even the Advanced Course is little more than refresher training).

A. Provide formal and standardized OJT along with tests which are conducted/monitored by an outside office or by a computerized testing system which reports results directly to ATC, MAJCOM's, or to the NDI Program Office.

<u>Solution Shows Promise</u>?

Not Promising				Very Promising
1	2	3	4	5
			3	3

Weighted Mean= 4.58
Standard Score= 60.5

<u>Solution is Attainable</u>?

Not Feasible				Very Feasible
1	2	3	4	5
		3	2	1

Weighted Mean= 3.83
Standard Score= 50.7

<u>Estimated Cost of Solution</u>?

Very High				Very Low
1	2	3	4	5
	2	3	1	

Weighted Mean= 2.88
Standard Score= 45.3

Promise + Feasibility = 8.41
Standard Score = 55.9

Promise + Feasibility + Cost = 11.29
Standard Score = 51.2

5.5 The Advanced Course is optional and does not introduce added depth of information.

A. Provide more structured training before and at the 7-level. The Advanced Course should be very advanced, so that each installation has one or more technician available for conducting training, conducting complex/novel inspections, and making final decisions on disputed calls. Insure that Advanced Course instructors are highly qualified and motivated.

Solution Shows Promise?

Not Promising				Very Promising	
1	2	3	4	5	
			1	5	

Weighted Mean= 4.85
Standard Score= 66.9

Solution is Attainable?

Not Feasible				Very Feasible	
1	2	3	4	5	
		1	4	1	*

Weighted Mean= 4.16
Standard Score= 56.5

Estimated Cost of Solution?

Very High				Very Low	
1	2	3	4	5	
	2	3		1	*

Weighted Mean= 3.28
Standard Score= 50.8

Promise + Feasibility = 9.01
Standard Score = 63.2

Promise + Feasibility + Cost = 12.29
Standard Score = 59.9

B. Conduct Advanced Course on-site by using a combination of on-site computer-based training and validation testing by remote computer-based testing or traveling instructors.

Solution Shows Promise?

Not Promising 1	2	3	4	Very Promising 5	
	1		4	1	*

Weighted Mean= 4.00
Standard Score= 46.7

Solution is Attainable?

Not Feasible 1	2	3	4	Very Feasible 5
	1	2	3	

Weighted Mean= 3.30
Standard Score= 41.4

Estimated Cost of Solution?

Very High 1	2	3	4	Very Low 5
	3	3		

Weighted Mean= 2.53
Standard Score= 40.5

Promise + Feasibility = 7.30
Standard Score = 42.3

Promise + Feasibility + Cost = 9.83
Standard Score = 38.5

C. Conduct Advanced Course and testing on-site by using traveling instructors for personnel unable to attend the Chanute course.

Solution Shows Promise?

Not Promising 1	2	3	4	Very Promising 5	
		1	5		*

Weighted Mean= 3.86
Standard Score= 43.3

Solution is Attainable?

Not Feasible 1	2	3	4	Very Feasible 5	
	1	4	1		*

Weighted Mean= 3.13
Standard Score= 38.4

Estimated Cost of Solution?

Very High 1	2	3	4	Very Low 5
	3	3		

Weighted Mean= 2.62
Standard Score= 41.8

Promise + Feasibility = 6.99
Standard Score = 38.5

Promise + Feasibility + Cost = 9.61
Standard Score = 36.6

5.6 There is no independent critical assessment of NDI proficiency (hands-on or theoretical knowledge) at critical milestones throughout the NDI career progression.

A. Develop and implement computer-based training/testing system with psychomotor/theory integrated training and TSD measurement built in. Test scenarios must be available from multiple sources: a) randomly generated (e.g., flaw location); b) ported in from external sources (tele-communications or portable memory); and c) resident in database of content questions. Test results can be automatically tele-communicated to ATC, the NDI Program Office, or to MAJCOM's to monitor quality control. Coordinate with the Training Systems Division of the Human Resources Laboratory and the Advanced Training System Branch at the Air Training Command to ensure that NDI considerations are made for automated trainers being developed.

<u>Solution Shows Promise</u>?

Not Promising				Very Promising
1	2	3	4	5
		1	2	3

Weighted Mean= 4.44
Standard Score= 57.1

<u>Solution is Attainable</u>?

Not Feasible				Very Feasible	
1	2	3	4	5	
	1	2	3		*

Weighted Mean= 3.37
Standard Score= 42.6

<u>Estimated Cost of Solution</u>?

Very High				Very Low
1	2	3	4	5
2	2	2		

Weighted Mean= 1.89
Standard Score= 31.7

Promise + Feasibility = 7.81
Standard Score = 48.5

Promise + Feasibility + Cost = 9.70
Standard Score = 37.4

B. Use traveling evaluators to periodically conduct impartial tests of technician proficiency.

Solution Shows Promise?

Not Promising 1	2	3	4	Very Promising 5
		3	2	1

Weighted Mean= 3.82
Standard Score= 42.4

Solution is Attainable?

Not Feasible 1	2	3	4	Very Feasible 5	
	3	2	1		*

Weighted Mean= 2.82
Standard Score= 33.0

Estimated Cost of Solution?

Very High 1	2	3	4	Very Low 5
	1	3	2	

Weighted Mean= 3.27
Standard Score= 50.7

Promise + Feasibility = 6.64
Standard Score = 34.2

Promise + Feasibility + Cost = 9.91
Standard Score = 39.2

5.7 There is occasionally a lack of updated training material and information on new testing equipment delivered to the field.

A. Develop and implement computer-based training/testing system that can be used to update training material and provide information on new testing equipment. Such a system should be designed so that updated information can be tele-communicated from a central location or so that portable memory units can be easily accessed.

Solution Shows Promise?

Not Promising				Very Promising
1	2	3	4	5
		1	2	3

Weighted Mean- 4.52
Standard Score- 50.0

Solution is Attainable?

Not Feasible				Very Feasible
1	2	3	4	5
		1	5	

Weighted Mean- 3.87
Standard Score- 51.4

Estimated Cost of Solution?

Very High				Very Low
1	2	3	4	5
1	3	2		

Weighted Mean- 2.20
Standard Score- 36.0

Promise + Feasibility - 8.39
Standard Score - 55.6

Promise + Feasibility + Cost - 10.59
Standard Score - 45.1

B. Implement MANPRINT-like procurement requirements to ensure that manpower/training/personnel issues are considered in the procurement process. Coordinate with the appropriate Air Force procurement agency to require that training/familiarization materials be provided by the manufacturer of any new testing equipment purchased (e.g., videotapes).

<u>Solution Shows Promise</u>?

Not Promising				Very Promising	
1	2	3	4	5	
			5	1	*

Weighted Mean- 4.22
Standard Score- 51.9

<u>Solution is Attainable</u>?

Not Feasible				Very Feasible	
1	2	3	4	5	
			4	2	*

Weighted Mean- 4.39
Standard Score- 60.5

<u>Estimated Cost of Solution</u>?

Very High				Very Low	
1	2	3	4	5	
	1	2	2	1	*

Weighted Mean- 3.68
Standard Score- 56.3

Promise + Feasibility - 8.61
Standard Score - 58.3

Promise + Feasibility + Cost - 12.29
Standard Score - 59.9

5.8 There is no vehicle to ensure that personnel get adequate practice using infrequent techniques.

A. Develop and implement computer-based training/testing system with psychomotor/theory integrated training and TSD measurement built in. Ensure that "practice modes" are available for all techniques, so that personnel can rehearse techniques without any performance scores being kept. Also, use computer-managed instruction to ensure that each technician is tested on infrequently used techniques at appropriate intervals. Coordinate with the Training Systems Division of the Human Resources Laboratory and the Advanced Training System Branch at the Air Training Command to ensure that NDI considerations are made for automated trainers being developed.

Solution Shows Promise?

Not Promising 1	2	3	4	Very Promising 5
		1	2	3

Weighted Mean= 4.36
Standard Score= 55.2

Solution is Attainable?

Not Feasible 1	2	3	4	Very Feasible 5
		2	4	

Weighted Mean= 3.64
Standard Score= 47.4

Estimated Cost of Solution?

Very High 1	2	3	4	Very Low 5
1	4	1		

Weighted Mean= 1.94
Standard Score= 32.4

Promise + Feasibility - 8.00
Standard Score - 50.9

Promise + Feasibility + Cost - 9.94
Standard Score - 39.5

B. Utilize a traveling instructor/evaluator to systematically conduct impartial tests of proficiency on all NDI techniques. This would motivate technicians to utilize existing training materials which address infrequent test procedures.

Solution Shows Promise?

Not Promising 1	2	3	4	Very Promising 5
	1	1	4	

Weighted Mean= 3.63
Standard Score= 37.9

Solution is Attainable?

Not Feasible 1	2	3	4	Very Feasible 5
	2	2	2	

Weighted Mean= 3.17
Standard Score= 39.1

Estimated Cost of Solution?

Very High 1	2	3	4	Very Low 5	
	1	3	1	1	*

Weighted Mean= 3.64
Standard Score= 55.8

Promise + Feasibility = 6.80
Standard Score = 36.2

Promise + Feasibility + Cost = 10.44
Standard Score = 43.8

6.0 Work Environment

6.1 NDI technicians and their supervisors should be made aware of the effects of environmental variables like temperature, noise, amount of light, humidity, extended time-on-task, and work schedule on human performance.

A. Develop and implement training information (e.g., videotape) to be used in initial training and provided to all supervisors in the field, that communicates what is known about the effects of environmental variables on NDI performance. In the material, provide suggestions about how such effects might be minimized.

<u>Solution Shows Promise</u>?

Not Promising 1	2	3	4	Very Promising 5
		3	1	2

Weighted Mean- 4.00
Standard Score- 46.7

<u>Solution is Attainable</u>?

Not Feasible 1	2	3	4	Very Feasible 5
		1	3	2

Weighted Mean- 4.20
Standard Score- 57.2

<u>Estimated Cost of Solution</u>?

Very High 1	2	3	4	Very Low 5	
	1		4	1	*

Weighted Mean- 4.00
Standard Score- 60.7

Promise + Feasibility - 8.20
Standard Score - 53.3

Promise + Feasibility + Cost - 12.20
Standard Score - 59.1

6.2 There is a lack of research on the effects of work and environmental variables on NDI performance.

A. Sponsor research on environmental variables that affect NDI technician performance. Coordinate such research with the Human Resources Laboratory, other Department of Defense NDI agencies, and other governmental/industrial agencies conducting NDI research (e.g., NRC and EPRI), to minimize redundancy. Incorporate existing knowledge of environmental factors into OJT and supervisor training modules. Establish guidelines based on what is known.

Solution Shows Promise?

Not Promising 1	2	3	4	Very Promising 5	
	4	2			*

Weighted Mean= 2.21
Standard Score= 4.0

Solution is Attainable?

Not Feasible 1	2	3	4	Very Feasible 5
		1	5	

Weighted Mean= 3.80
Standard Score= 50.2

Estimated Cost of Solution?

Very High 1	2	3	4	Very Low 5	
	1	3	1	1	*

Weighted Mean= 3.41
Standard Score= 52.6

Promise + Feasibility = 6.01
Standard Score = 26.5

Promise + Feasibility + Cost = 9.42
Standard Score = 35.0

7.0 Performance

7.1 Because of the nature of NDI tasks, there is an intrinsic lack of performance feedback. For example, there is obviously no real-world immediate feedback for missed calls, and often delayed or no feedback for hits and false alarms. Implications extend to assessment, training, motivation, accountability, and self-confidence.

A. Develop and implement computer-based training/testing system with psychomotor/theory integrated training and TSD measurement built in to provide immediate and meaningful feedback about correct calls, missed calls, false calls, correct no-calls, sensitivity, and response bias. This feedback about proficiency should be available to students, OJT instructors, ATC, NDI Program Office, MAJCOM's, and depots. Coordinate with the Training Systems Division of the Human Resources Laboratory and the Advanced Training System Branch at the Air Training Command to ensure that NDI considerations are made for automated trainers being developed.

Solution Shows Promise?

Not Promising				Very Promising
1	2	3	4	5
		1	2	3

Weighted Mean= 4.42
Standard Score= 56.7

Solution is Attainable?

Not Feasible				Very Feasible
1	2	3	4	5
	1		5	*

Weighted Mean= 3.78
Standard Score= 49.8

Estimated Cost of Solution?

Very High				Very Low
1	2	3	4	5
1	3	2		

Weighted Mean= 2.06
Standard Score= 34.0

Promise + Feasibility = 8.20
Standard Score = 53.3

Promise + Feasibility + Cost = 10.26
Standard Score = 42.3

B. Within the framework of the existing system, formalize and reinforce a checking system so that calls are checked by at least two other (higher-level) technicians. Missed calls are more difficult to measure but different strategies should be used periodically to keep the technicians motivated (especially inexperienced technicians): a) conduct spot-check inspections for quality control, and b) to the extent that time allows, have two technicians inspect same test piece.

Solution Shows Promise?

Not Promising				Very Promising	
1	2	3	4	5	
	1		4	1	*

Weighted Mean= 3.86
Standard Score= 43.3

Solution is Attainable?

Not Feasible				Very Feasible	
1	2	3	4	5	
1		3	2		*

Weighted Mean= 2.89
Standard Score= 34.2

Estimated Cost of Solution?

Very High				Very Low	
1	2	3	4	5	
	2	2	1	1	*

Weighted Mean= 3.27
Standard Score= 50.7

Promise + Feasibility = 6.75
Standard Score = 35.6

Promise + Feasibility + Cost = 10.02
Standard Score = 40.2

C. Provide more example test pieces with and without known flaws. Require technicians to periodically inspect these pieces and provide immediate performance feedback. Confirm practice by remote testing with computer-based testing or with traveling evaluator.

Solution Shows Promise?

Not Promising				Very Promising	
1	2	3	4	5	
			3	3	

Weighted Mean= 4.54
Standard Score= 59.5

Solution is Attainable?

Not Feasible				Very Feasible	
1	2	3	4	5	
			3	3	

Weighted Mean= 4.57
Standard Score= 63.7

Estimated Cost of Solution?

Very High				Very Low	
1	2	3	4	5	
	3	2		1	*

Weighted Mean= 3.00
Standard Score= 47.0

Promise + Feasibility = 9.11
Standard Score = 64.4

Promise + Feasibility + Cost = 12.11
Standard Score = 58.3

7.2 There is a lack of systematic research in the area of NDI performance. Also, what is known from other fields is not often applied or communicated to technicians or managers.

A. Sponsor more applied research in this field. Coordinate with the Air Force Human Resources Laboratory to exploit more about what is known about vigilance and training technology in this area. The literature offers many promising research topics. Provide these topics (as presented in the Areas of Concern and Possible Solutions sections of this report) to participating researchers.

Solution Shows Promise?

Not Promising 1	2	3	4	Very Promising 5
		2	3	1

Weighted Mean= 3.95
Standard Score= 45.5

Solution is Attainable?

Not Feasible 1	2	3	4	Very Feasible 5
		1	4	1

Weighted Mean= 4.05
Standard Score= 54.6

Estimated Cost of Solution?

Very High 1	2	3	4	Very Low 5
	2	3	1	

Weighted Mean= 3.00
Standard Score= 47.0

Promise + Feasibility = 8.00
Standard Score = 50.9

Promise + Feasibility + Cost = 11.00
Standard Score = 48.7

B. Publish more technical information in the Air Force NDI newsletter and subscribe to other technical journals for each laboratory.

<u>Solution Shows Promise</u>?

Not Promising				Very Promising
1	2	3	4	5
			6	*****

Weighted Mean= 4.00
Standard Score= 46.7

<u>Solution is Attainable</u>?

Not Feasible				Very Feasible
1	2	3	4	5
		1	**3**	**2**

Weighted Mean= 4.14
Standard Score= 56.1

<u>Estimated Cost of Solution</u>?

Very High				Very Low
1	2	3	4	5
		1	**2**	**3**

Weighted Mean= 4.40
Standard Score= 66.2

Promise + Feasibility = 8.14
Standard Score = 52.6

Promise + Feasibility + Cost = 12.54
Standard Score = 62.1

7.3 NDI tasks and procedures are often not designed with human performance in mind.

A. Because the Air Force system relies heavily upon the appropriate execution of established procedures, a strict quality-control assessment should be made of all new procedures and the current updating/correcting system for existing procedures must be maintained/ increased. Procurement systems for both NDI test equipment and new weapon systems should be made aware of the MANPRINT-like considerations which should be made for new equipment requiring NDI.

<u>Solution Shows Promise</u>?

Not Promising 1	2	3	4	Very Promising 5	
		1	4	1	*

Weighted Mean= 4.05
Standard Score= 47.9

<u>Solution is Attainable</u>?

Not Feasible 1	2	3	4	Very Feasible 5	
		3	3		*

Weighted Mean= 3.60
Standard Score= 46.7

<u>Estimated Cost of Solution</u>?

Very High 1	2	3	4	Very Low 5
	1	3	2	

Weighted Mean= 3.36
Standard Score= 51.9

Promise + Feasibility = 7.65
Standard Score = 46.6

Promise + Feasibility + Cost = 11.01
Standard Score = 48.8

7.4 Nondestructive testing performance is not adequate.

(Obviously, all suggested solutions in this report are targeted at this concern; the following candidate solutions address specific performance issues.)

A. In order to increase the level of detection on very critical or difficult tests, use redundant inspections (e.g., different technicians apply the same technique to a test piece, multiple techniques are applied to a test piece, or different technicians simultaneously apply a test).

Solution Shows Promise?

Not Promising				Very Promising	
1	2	3	4	5	
		2	3	1	

Weighted Mean= 3.85
Standard Score= 43.1

Solution is Attainable?

Not Feasible				Very Feasible	
1	2	3	4	5	
	1	2	3		*

Weighted Mean= 3.35
Standard Score= 42.3

Estimated Cost of Solution?

Very High				Very Low	
1	2	3	4	5	
1		2	2	1	*

Weighted Mean= 3.24
Standard Score= 50.3

Promise + Feasibility = 7.20
Standard Score = 41.1

Promise + Feasibility + Cost = 10.44
Standard Score = 43.8

B. Conduct human reliability analyses on each technique, feed results back to the training and equipment procurement communities, and use results to develop portable performance aids for technicians in the field.

<u>Solution Shows Promise</u>?

Not Promising				Very Promising	
1	2	3	4	5	
		1	5		*

Weighted Mean= 3.86
Standard Score= 43.3

<u>Solution is Attainable</u>?

Not Feasible				Very Feasible
1	2	3	4	5
		1	3	2

Weighted Mean= 4.21
Standard Score= 57.4

<u>Estimated Cost of Solution</u>?

Very High				Very Low	
1	2	3	4	5	
	1	3	1	1	*

Weighted Mean= 3.54
Standard Score= 54.4

Promise + Feasibility = 8.07
Standard Score = 51.7

Promise + Feasibility + Cost = 11.61
Standard Score = 54.0

C. Create and maintain an electronic database that has at least three cross-referenced components: a) historical and theoretical (e.g., from fracture mechanics) information about location, orientation, type, etc., of detected flaws for a specific component; b) human-reliability information about what human errors are most common for a specific procedure/ technique; and c) information about the content of the current set of TO's. This database could be used by NDI training/proficiency evaluators, equipment developers, and TO developers/evaluators.

<u>Solution Shows Promise</u>?

Not Promising				Very Promising	
1	2	3	4	5	
			4	2	*

Weighted Mean= 4.36
Standard Score= 55.2

<u>Solution is Attainable</u>?

Not Feasible				Very Feasible
1	2	3	4	5
		3	2	1

Weighted Mean= 3.84
Standard Score= 50.9

<u>Estimated Cost of Solution</u>?

Very High				Very Low
1	2	3	4	5
	3	2	1	

Weighted Mean= 2.73
Standard Score= 43.3

Promise + Feasibility = 8.20
Standard Score = 53.3

Promise + Feasibility + Cost = 10.93
Standard Score = 48.1

7.5 NDI personnel report barriers to practicing some techniques at some installations.

A. Use automated trainer/tester to ensure that practice opportunities exist for all methods at all locations. Use tele-communications or traveling evaluator to confirm that adequate sustainment practice has been conducted by all technicians on all methods.

Solution Shows Promise?

Not Promising 1	2	3	4	Very Promising 5
		1	4	1

Weighted Mean= 4.04
Standard Score= 47.6

Solution is Attainable?

Not Feasible 1	2	3	4	Very Feasible 5
		1	4	1

Weighted Mean= 4.05
Standard Score= 54.6

Estimated Cost of Solution?

Very High 1	2	3	4	Very Low 5
	3	3		

Weighted Mean= 2.60
Standard Score= 41.5

Promise + Feasibility = 8.09
Standard Score = 52.0

Promise + Feasibility + Cost = 10.69
Standard Score = 46.0

B. Reconfirm the presence of training materials and an adequate number of test pieces for each technique at each laboratory location. Use existing MAJCOM/MSEP inspection teams to confirm that adequate sustainment practice has been conducted by all technicians on all techniques.

<u>Solution Shows Promise</u>?

Not Promising 1	2	3	4	Very Promising 5	
		1	5		

Weighted Mean= 3.82
Standard Score= 42.4

<u>Solution is Attainable</u>?

Not Feasible 1	2	3	4	Very Feasible 5	
			4	2	*

Weighted Mean= 4.35
Standard Score= 59.8

<u>Estimated Cost of Solution</u>?

Very High 1	2	3	4	Very Low 5	
		4	1	1	

Weighted Mean= 3.64
Standard Score= 55.8

Promise + Feasibility = 8.17
Standard Score = 52.9

Promise + Feasibility + Cost = 11.81
Standard Score = 55.7

7.6 If the operator approach is taken, then temptations must be resisted to increase demands on the operators.

A. Even though the technician has been identified as the goal of the NDI program, the system still currently relies heavily upon the integrity of the TO, especially with new NDI technicians. The Air Force should ensure: a) that TO's maintain extremely high quality and are completely understandable; b) that equipment is standardized, is user-friendly, and conforms to current TO's; and c) that job demands are not placed on young technicians which assume technician-level knowledge/skills.

Solution Shows Promise?

Not Promising 1	2	3	4	Very Promising 5
		1	3	2

Weighted Mean- 4.40
Standard Score- 56.2

Solution is Attainable?

Not Feasible 1	2	3	4	Very Feasible 5
	2	3	1	

Weighted Mean- 3.00
Standard Score- 36.2

Estimated Cost of Solution?

Very High 1	2	3	4	Very Low 5
	3	3		

Weighted Mean- 2.50
Standard Score- 40.1

Promise + Feasibility - 7.40
Standard Score - 43.5

Promise + Feasibility + Cost - 9.90
Standard Score - 39.1

8.0 Procedures

8.1 Most assessments of NDI proficiency have focussed on the detection task. While this is the most important task, there are other human tasks involved. Human error can occur in any tasks and reduce proficiency. No human reliability assessments could be found in which error rates are assessed for all tasks/subtasks. Consequently, even if the existing test results were valid and were perfect indicators of actual field performance, there would still be no true assessment of actual NDI proficiency.

A. Conduct human reliability analysis on each technique, feed results back to the training, TO development, and equipment procurement communities, and develop/exploit portable performance aids for technicians in the field.

Solution Shows Promise?

Not Promising				Very Promising
1	2	3	4	5
		2	4	

Weighted Mean= 3.75
Standard Score= 40.7

Solution is Attainable?

Not Feasible				Very Feasible
1	2	3	4	5
		2	4	

Weighted Mean= 3.71
Standard Score= 48.6

Estimated Cost of Solution?

Very High				Very Low
1	2	3	4	5
		6		

Weighted Mean= 3.00
Standard Score= 47.0

Promise + Feasibility = 7.46
Standard Score = 44.2

Promise + Feasibility + Cost = 10.46
Standard Score = 44.0

B. Create and maintain an electronic database that has at least three cross-referenced components: a) historical and theoretical (e.g., from fracture mechanics) information about location, orientation, type, etc., of detected flaws for a specific component; b) human-reliability information about what human errors are most common for a specific procedure/technique; and c) information about the content of the current set of TO's. This database could be used by NDI training/proficiency evaluators, equipment developers, and TO developers/evaluators.

Solution Shows Promise?

Not Promising 1	2	3	4	Very Promising 5
		2	2	2

Weighted Mean= 4.10
Standard Score= 49.0

Solution is Attainable?

Not Feasible 1	2	3	4	Very Feasible 5
		1	5	

Weighted Mean= 3.83
Standard Score= 50.7

Estimated Cost of Solution?

Very High 1	2	3	4	Very Low 5
	1	3	2	

Weighted Mean= 3.27
Standard Score= 50.7

Promise + Feasibility = 7.93
Standard Score = 50.0

Promise + Feasibility + Cost = 11.20
Standard Score = 50.4

8.2 To the extent that the Air Force embraces the operator as opposed to the technician approach to NDI, there are implications for the procedures used. Extensive quality control procedures for technical orders must be strictly followed and tested to ensure that there is no possibility of misinterpretation. In addition, steps should be taken to minimize possible negative side-effects such as boredom.

A. Because new NDI personnel are essentially operators, not technicians, it is absolutely critical that technique development of TO's be high in quality, understandable, and anticipate most possible variations that could reasonably be expected in the field. During TO development, actual technicians should be used to test and validate/verify TO's. The existing system for updating TO's from the field appears to be excellent, but could become more formalized by the formation of an electronic database.

Solution Shows Promise?

Not Promising 1	2	3	4	Very Promising 5
			2	4

Weighted Mean= 4.14
Standard Score= 50.0

Solution is Attainable?

Not Feasible 1	2	3	4	Very Feasible 5
			6	*

Weighted Mean= 4.00
Standard Score= 53.7

Estimated Cost of Solution?

Very High 1	2	3	4	Very Low 5
	2	1	3	

Weighted Mean= 3.33
Standard Score= 51.5

Promise + Feasibility = 8.14
Standard Score = 52.6

Promise + Feasibility + Cost = 11.47
Standard Score = 52.8

B. The criticality of parts should be determined, identified in the TO, and possibly identified on the actual part. Special attention should be devoted to make sure that procedures for inspecting highly critical parts are clear, specific, and detailed. Procedures should be proofed by a sample of representative AF technicians who will be conducting such tests in the field.

<u>Solution Shows Promise</u>?

Not Promising 1	2	3	4	Very Promising 5	
			5	1	*

Weighted Mean= 4.21
Standard Score= 51.7

<u>Solution is Attainable</u>?

Not Feasible 1	2	3	4	Very Feasible 5
		2	3	1

Weighted Mean= 3.92
Standard Score= 52.3

<u>Estimated Cost of Solution</u>?

Very High 1	2	3	4	Very Low 5
	2	2	2	

Weighted Mean= 3.12
Standard Score= 48.6

Promise + Feasibility = 8.13
Standard Score = 52.4

Promise + Feasibility + Cost = 11.25
Standard Score = 50.9

9.0 Motivation

9.1 The lack of feedback that is characteristic of NDI tasks affects the motivation of the NDI personnel.

A. Develop and implement computer-based training/testing system with psychomotor/theory integrated training and TSD measurement built in to provide immediate and meaningful feedback for students and OJT instructors. Include competitive programs where technicians can compete with a computerized "average technician," their own personal score on past exercises, or with other technicians at their laboratory. Coordinate with the Training Systems Division of the Human Resources Laboratory and the Advanced Training System Branch at the Air Training Command to ensure that NDI considerations are made for automated trainers being developed.

<u>Solution Shows Promise</u>?

Not Promising				Very Promising
1	2	3	4	5
		1	2	3

Weighted Mean= 4.39
Standard Score= 56.0

<u>Solution is Attainable</u>?

Not Feasible				Very Feasible
1	2	3	4	5
		1	5	

Weighted Mean= 3.88
Standard Score= 51.6

<u>Estimated Cost of Solution</u>?

Very High				Very Low
1	2	3	4	5
	4	2		

Weighted Mean= 2.40
Standard Score= 38.7

Promise + Feasibility = 8.27
Standard Score = 54.2

Promise + Feasibility + Cost = 10.67
Standard Score = 45.8

B. Within the framework of the existing system, formalize and reinforce a checking system so that calls are checked by at least two other (higher-level) technicians. Missed calls are more difficult to measure but different strategies should be used periodically to keep the technicians motivated (especially inexperienced technicians): a) conduct spot-check inspections for quality control and b) to the extent that time allows, have two technicians inspect same test piece.

<u>Solution Shows Promise</u>?

Not Promising				Very Promising	
1	2	3	4	5	
	1		5		*

Weighted Mean= 3.64
Standard Score= 38.1

<u>Solution is Attainable</u>?

Not Feasible				Very Feasible
1	2	3	4	5
	3		3	

Weighted Mean= 2.90
Standard Score= 34.4

<u>Estimated Cost of Solution</u>?

Very High				Very Low	
1	2	3	4	5	
1	2	1	1	1	*

Weighted Mean= 2.80
Standard Score= 44.2

Promise + Feasibility = 6.54
Standard Score = 33.0

Promise + Feasibility + Cost = 9.34
Standard Score = 34.3

C. Display a laboratory graph or sign depicting estimated "lives saved" or "money saved" due to cost avoidance because of NDI flaws detected.

Solution Shows Promise?

Not Promising 1	2	3	4	Very Promising 5	
		2	4		*

Weighted Mean= 3.67
Standard Score= 38.8

Solution is Attainable?

Not Feasible 1	2	3	4	Very Feasible 5	
2	4				*

Weighted Mean= 1.60
Standard Score= 11.6

Estimated Cost of Solution?

Very High 1	2	3	4	Very Low 5
			2	4

Weighted Mean= 4.70
Standard Score= 70.3

Promise + Feasibility = 5.27
Standard Score = 17.4

Promise + Feasibility + Cost = 9.97
Standard Score = 39.7

9.2 The effects of response bias and the variables that influence response bias have been ignored in most research on NDI performance proficiency.

A. Both technicians and supervisors must be educated in the basic fundamentals of response bias and sensitivity and their effects on performance in a detection task. Also, variables known to affect response bias should be explained. A realistic (and attainable) hit/false alarm rate should be determined and, for some tests which are especially high or low in criticality, special instructions should be provided to establish appropriate hit/false-alarm rates.

Solution Shows Promise?

Not Promising 1	2	3	4	Very Promising 5	
	1		4	1	*

Weighted Mean= 3.86
Standard Score= 43.3

Solution is Attainable?

Not Feasible 1	2	3	4	Very Feasible 5
		3	2	1

Weighted Mean= 3.80
Standard Score= 50.2

Estimated Cost of Solution?

Very High 1	2	3	4	Very Low 5
		3	3	

Weighted Mean= 3.69
Standard Score= 56.5

Promise + Feasibility = 7.66
Standard Score = 46.7

Promise + Feasibility + Cost = 11.35
Standard Score = 51.7

9.3 There is a reported lack of confidence and self-efficacy among nondestructive testing personnel.

A. Personnel cannot be high in self efficacy if they receive no feedback about their proficiency. The largest gains in self-efficacy will result by making changes suggested in some of the earlier sections: a) improve initial training and OJT, b) adopt a defendable, standardized, meaningful performance measurement/ testing system, and c) provide ample practice with feedback about proficiency.

<u>Solution Shows Promise</u>?

Not Promising 1	2	3	4	Very Promising 5
			3	3

Weighted Mean= 4.62
Standard Score= 61.4

<u>Solution is Attainable</u>?

Not Feasible 1	2	3	4	Very Feasible 5
		1	3	2

Weighted Mean= 4.33
Standard Score= 59.5

<u>Estimated Cost of Solution</u>?

Very High 1	2	3	4	Very Low 5
		4	2	

Weighted Mean= 3.47
Standard Score= 53.4

Promise + Feasibility = 8.95
Standard Score = 62.5

Promise + Feasibility + Cost = 12.42
Standard Score = 61.0

9.4 The professionalism, feelings of importance in the maintenance community, and general esprit de corps could be improved for NDI.

A. Create a plan to promote the importance, goals, tasks, and professionalism of NDI (e.g., videotapes, pamphlets, open houses, conduct conferences, promotional signs, etc.).

<u>Solution Shows Promise</u>?

Not Promising 1	2	3	4	Very Promising 5
		1	3	2

Weighted Mean= 4.30
Standard Score= 53.8

<u>Solution is Attainable</u>?

Not Feasible 1	2	3	4	Very Feasible 5	
	1		4	1	*

Weighted Mean= 3.82
Standard Score= 50.5

<u>Estimated Cost of Solution</u>?

Very High 1	2	3	4	Very Low 5
			2	4

Weighted Mean= 4.74
Standard Score= 70.9

Promise + Feasibility = 8.12
Standard Score = 52.3

Promise + Feasibility + Cost = 12.86
Standard Score = 64.9

9.5 There are reports of low motivation among technicians in the nondestructive testing community.

A. Some NDI tasks might intrinsically produce low motivation. However, the effects of monotony can be reduced by providing adequate work conditions, proper work schedule, increased training, and immediate/ accurate performance feedback. Ironically, some human performance tasks with similar characteristics (looking and waiting for an unusual event) are known to dramatically increase motivation, as long as some minimal amount of feedback is occasionally provided (e.g., gold prospecting, gambling, organisms on partial reinforcement schedules, etc.). However, some sort of monetary reward is usually involved in each of these examples. While it seems unfeasible to offer monetary reward for correct decisions, providing feedback about performance, graphing a person's performance, conducting competition among colleagues, and public acknowledgement of superior performance are examples of how meaningful non-monetary rewards might be used. Also, any steps taken to improve training, self-efficacy, and esprit will also probably increase general motivation.

<u>Solution Shows Promise</u>?

Not Promising				Very Promising
1	2	3	4	5
			5	1

Weighted Mean= 4.23
Standard Score= 52.1

<u>Solution is Attainable</u>?

Not Feasible				Very Feasible
1	2	3	4	5
		2	2	2

Weighted Mean= 4.20
Standard Score= 57.2

<u>Estimated Cost of Solution</u>?

Very High				Very Low
1	2	3	4	5
			2	4

Weighted Mean= 4.78
Standard Score= 71.4

Promise + Feasibility = 8.43
Standard Score = 56.1

Promise + Feasibility + Cost = 13.21
Standard Score = 67.9

10.0 System

10.1 There is a heterogeneity among personnel and NDI tasks across different units, which makes it more difficult to prescribe general, standardized remedies.

A. Standardization of OJT and adopting a standard measurement system (TSD) should help standardize tasks and expectations. Also, implementation of standardized (computer-based or other) practice opportunities and performance tests, whether conducted remotely with tele-communications or by a traveling evaluator, will facilitate standardization.

Solution Shows Promise?

Not Promising 1	2	3	4	Very Promising 5
		1	4	1

Weighted Mean= 4.05
Standard Score= 47.9

Solution is Attainable?

Not Feasible 1	2	3	4	Very Feasible 5
		1	4	1

Weighted Mean= 4.05
Standard Score= 54.6

Estimated Cost of Solution?

Very High 1	2	3	4	Very Low 5
		5	1	

Weighted Mean= 3.23
Standard Score= 50.1

Promise + Feasibility = 8.10
Standard Score = 52.1

Promise + Feasibility + Cost = 11.33
Standard Score = 51.6

10.2 There needs to be more opportunity for practice, especially for infrequently used procedures.

A. Use automated trainer/tester to ensure that practice opportunities exist for all methods at all locations. Use tele-communications or traveling evaluator to confirm that adequate sustainment practice has been conducted by all technicians on all methods.

Solution Shows Promise?

Not Promising 1	2	3	4	Very Promising 5
			4	2

Weighted Mean= 4.43
Standard Score= 56.9

Solution is Attainable?

Not Feasible 1	2	3	4	Very Feasible 5
	1	1	4	

Weighted Mean= 3.53
Standard Score= 45.4

Estimated Cost of Solution?

Very High 1	2	3	4	Very Low 5
	1	5		

Weighted Mean= 2.79
Standard Score= 44.1

Promise + Feasibility = 7.96
Standard Score = 50.4

Promise + Feasibility + Cost = 10.75
Standard Score = 46.5

B. Reconfirm the presence of training materials and an adequate number of test pieces for each technique at each laboratory location. Use existing MAJCOM inspection teams to confirm that adequate sustainment practice has been conducted by all technicians on all techniques.

<u>Solution Shows Promise</u>?

Not Promising				Very Promising
1	2	3	4	5
		2	1	3

Weighted Mean= 4.24
Standard Score= 52.4

<u>Solution is Attainable</u>?

Not Feasible				Very Feasible
1	2	3	4	5
		3	1	2

Weighted Mean= 3.95
Standard Score= 52.8

<u>Estimated Cost of Solution</u>?

Very High				Very Low
1	2	3	4	5
	2	4		1

Weighted Mean= 3.00
Standard Score= 47.0

Promise + Feasibility - 8.19
Standard Score - 53.2

Promise + Feasibility + Cost - 11.19
Standard Score - 50.3

10.3 Because of the lack of adequate measures of performance, standards and qualifications for career advancement are less defendable.

A. Adopting a standardized measurement system, providing accurate feedback to technicians and their supervisors, forming a database on technician performance on different NDI tasks, and conducting periodic validated tests (either with a computerized test system or with a traveling evaluator) will help provide the solution to this concern. Quantified, validated, reliable, and normalized measures of technician performance will provide an additional tool to help make evaluation, work assignments, supplemental training, and other personnel decisions much easier and defendable for supervisors, and understandable for the technicians.

<u>Solution Shows Promise</u>?

Not Promising				Very Promising	
1	2	3	4	5	
			6		*

Weighted Mean= 4.00
Standard Score= 46.7

<u>Solution is Attainable</u>?

Not Feasible				Very Feasible	
1	2	3	4	5	
		3	3		*

Weighted Mean= 3.53
Standard Score= 45.4

<u>Estimated Cost of Solution</u>?

Very High				Very Low	
1	2	3	4	5	
	1	2	2	1	*

Weighted Mean= 3.73
Standard Score= 57.0

Promise + Feasibility = 7.53
Standard Score = 45.1

Promise + Feasibility + Cost = 11.26
Standard Score = 51.0

11.0 Measurement

11.1 The measures of proficiency used in the past by researchers investigating NDI technician performance are not good measures of true detection ability (sensitivity).

A. The Air Force should adopt TSD methodology as a standard measurement system. This involves educating the NDI community about the general principles of TSD and what the two measures mean.

<u>Solution Shows Promise</u>?

Not Promising				Very Promising
1	2	3	4	5
			3	3

Weighted Mean= 4.60
Standard Score= 61.0

<u>Solution is Attainable</u>?

Not Feasible				Very Feasible
1	2	3	4	5
			4	2

Weighted Mean= 4.45
Standard Score= 61.6

<u>Estimated Cost of Solution</u>?

Very High				Very Low
1	2	3	4	5
		3	3	

Weighted Mean= 3.60
Standard Score= 55.2

Promise + Feasibility = 9.05
Standard Score = 63.7

Promise + Feasibility + Cost = 12.65
Standard Score = 63.0

11.2 The measures of proficiency used in the past by researchers investigating NDI technician performance are not good measures of the technician's response bias.

A. Incorporate signal detection measures of response bias into assessment procedures. The general logic of response bias should be taught to all technicians and incorporated into any NDI automated trainers. Each TO should specify the response bias appropriate for a given test piece.

Solution Shows Promise?

Not Promising 1	2	3	4	Very Promising 5
		1	3	2

Weighted Mean= 4.28
Standard Score= 53.3

Solution is Attainable?

Not Feasible 1	2	3	4	Very Feasible 5
		3	3	

Weighted Mean= 3.53
Standard Score= 45.4

Estimated Cost of Solution?

Very High 1	2	3	4	Very Low 5
		5		1

Weighted Mean= 3.46
Standard Score= 53.3

Promise + Feasibility = 7.81
Standard Score = 48.5

Promise + Feasibility + Cost = 11.27
Standard Score = 51.0

11.3 The experimental design or approach taken in the past by researchers investigating NDI technician performance were sometimes inappropriate or led to misinterpretation.

A. For each technique, a generic test protocol should be developed and implemented in all future research investigating NDI technician proficiency. The protocols should take into account and attempt to control for variables from the human experimentation literature which are known to contaminate results. The protocols should demand TSD measurement, appropriate sample size(s), and appropriate statistical analyses. They should attempt to address all criticisms of previous test protocols in the NDI proficiency literature.

<u>Solution Shows Promise</u>?

Not Promising 1	2	3	4	Very Promising 5	
		2	3	1	*

Weighted Mean= 3.86
Standard Score= 43.3

<u>Solution is Attainable</u>?

Not Feasible 1	2	3	4	Very Feasible 5
		1	3	2

Weighted Mean= 4.28
Standard Score= 58.6

<u>Estimated Cost of Solution</u>?

Very High 1	2	3	4	Very Low 5
		4	1	1

Weighted Mean= 3.40
Standard Score= 52.5

Promise + Feasibility = 8.14
Standard Score = 52.6

Promise + Feasibility + Cost = 11.54
Standard Score = 53.4

11.4 Because of inappropriate measures of proficiency, there might be inaccuracies in the literature with regard to potential correlates of NDI proficiency.

A. All conclusions based on earlier inadequate measures of proficiency and how they relate to selection variables and training variables must be questioned. If possible, re-analysis should be conducted. Any new research should take these considerations into account.

<u>Solution Shows Promise</u>?

Not Promising 1	2	3	4	Very Promising 5	
	1		4	1	*

Weighted Mean= 4.00
Standard Score= 46.7

<u>Solution is Attainable</u>?

Not Feasible 1	2	3	4	Very Feasible 5
		1	3	2

Weighted Mean= 4.16
Standard Score= 56.5

<u>Estimated Cost of Solution</u>?

Very High 1	2	3	4	Very Low 5	
	1	3	1	1	*

Weighted Mean= 3.21
Standard Score= 49.9

Promise + Feasibility = 8.16
Standard Score = 52.8

Promise + Feasibility + Cost = 11.37
Standard Score = 51.9

11.5 The lack of an appropriate measure of proficiency has potential wide-ranging negative effects ranging from difficulty in training to low motivation/morale.

A. Adopting TSD measures of sensitivity and response bias and educating the NDI community about those measures should help solve this concern.

Solution Shows Promise?

Not Promising 1	2	3	4	Very Promising 5
			4	2

Weighted Mean= 4.47
Standard Score= 57.9

Solution is Attainable?

Not Feasible 1	2	3	4	Very Feasible 5
			4	2

Weighted Mean= 4.35
Standard Score= 59.8

Estimated Cost of Solution?

Very High 1	2	3	4	Very Low 5
		2	3	1

Weighted Mean= 3.76
Standard Score= 57.4

Promise + Feasibility = 8.82
Standard Score = 60.9

Promise + Feasibility + Cost = 12.58
Standard Score = 62.4

12.0 Equipment

12.1 Equipment is not currently automated.

A. Equipment should be selected which is semi-automated and reduces the number and complexity of the decisions made by the inspector (especially for new technicians who do not thoroughly understand the theory behind NDI testing).

Solution Shows Promise?

Not Promising 1	2	3	4	Very Promising 5	
			6		*

Weighted Mean= 4.00
Standard Score= 46.7

Solution is Attainable?

Not Feasible 1	2	3	4	Very Feasible 5	
	1	3	2		*

Weighted Mean= 3.11
Standard Score= 43.2

Estimated Cost of Solution?

Very High 1	2	3	4	Very Low 5	
2	2	1	1		*

Weighted Mean= 2.11
Standard Score= 34.8

Promise + Feasibility = 7.11
Standard Score = 40.0

Promise + Feasibility + Cost = 9.22
Standard Score = 33.2

12.2 NDI equipment (displays, controls, and sensors) is not adequately standardized.

A. NDI equipment (displays, controls, and sensors) should be standardized as much as possible. Steps should be taken to ensure that the equipment procurement process considers standardization. Coordination should be made with Air Force Materials Lab, nondestructive evaluation technical societies, and the National Bureau of Standards (agencies which have already initiated steps to begin the standardization process for NDI equipment).

Solution Shows Promise?

Not Promising				Very Promising	
1	2	3	4	5	
	1		5		*

Weighted Mean= 3.64
Standard Score= 38.1

Solution is Attainable?

Not Feasible				Very Feasible
1	2	3	4	5
		2	3	1

Weighted Mean= 4.00
Standard Score= 53.7

Estimated Cost of Solution?

Very High				Very Low	
1	2	3	4	5	
	1	2	2	1	*

Weighted Mean= 3.56
Standard Score= 54.7

Promise + Feasibility = 7.64
Standard Score = 46.4

Promise + Feasibility + Cost = 11.20
Standard Score = 50.4

12.3 NDI equipment displays, controls, and sensors have not always been designed, evaluated, and selected on the basis of their ease of use among actual NDI technicians.

A. Human-engineering guidelines should be developed and provided to (demanded of) equipment manufacturers. Where possible, empirically founded recommendations in the existing literature should be included. Where knowledge voids exist, sponsor human-factors research using AF NDI technicians as subjects to determine optimal design features.

Solution Shows Promise?

Not Promising 1	2	3	4	Very Promising 5	
		1	5		*

Weighted Mean= 3.79
Standard Score= 41.7

Solution is Attainable?

Not Feasible 1	2	3	4	Very Feasible 5
	1	3	2	

Weighted Mean= 3.06
Standard Score= 37.2

Estimated Cost of Solution?

Very High 1	2	3	4	Very Low 5
	4		2	

Weighted Mean= 2.62
Standard Score= 41.8

Promise + Feasibility = 6.85
Standard Score = 36.8

Promise + Feasibility + Cost = 9.47
Standard Score = 35.4

12.4 Because of the criticality of the technical orders in Air Force NDI, great care must be taken to ensure that all technical orders correspond to current equipment available.

A. Even with a technician as the ultimate goal, there will still be many new personnel who are rotely following TO's. Steps must be taken to ensure that whenever a new TO is written, it can be conducted with all existing NDI equipment in the field. Similarly, whenever a new piece of equipment is adopted, feed-forward information should be provided so that TO's can be appropriately modified.

<u>Solution Shows Promise</u>?

Not Promising 1	2	3	4	Very Promising 5
		1	4	1

Weighted Mean= 4.14
Standard Score= 50.0

<u>Solution is Attainable</u>?

Not Feasible 1	2	3	4	Very Feasible 5
		2	2	2

Weighted Mean= 4.11
Standard Score= 55.6

<u>Estimated Cost of Solution</u>?

Very High 1	2	3	4	Very Low 5
		3	3	

Weighted Mean= 3.50
Standard Score= 53.8

Promise + Feasibility = 8.25
Standard Score = 53.9

Promise + Feasibility + Cost = 11.75
Standard Score = 55.2

12.5 There is a lack of knowledge and research in the areas of display design, controls design, and sensor design to maximize human performance.

A. Sponsor research in human-factors design of NDI equipment and exploit the findings of other agencies conducting similar research (e.g., NRC and EPRI).

Solution Shows Promise?

Not Promising 1	2	3	4	Very Promising 5
		2	3	1

Weighted Mean= 3.86
Standard Score= 43.3

Solution is Attainable?

Not Feasible 1	2	3	4	Very Feasible 5
		4	2	

Weighted Mean= 3.22
Standard Score= 40.0

Estimated Cost of Solution?

Very High 1	2	3	4	Very Low 5	
	1	4		1	*

Weighted Mean= 3.33
Standard Score= 51.5

Promise + Feasibility = 7.08
Standard Score = 39.6

Promise + Feasibility + Cost = 10.41
Standard Score = 43.6

12.6 Weapon-system manufacturers do not always consider NDI testing when designing a new system.

A. Coordinate with the Air Force Materials Lab, nondestructive evaluation technical societies, and National Bureau of Standards which have already taken steps to begin the standardization process. Ensure that appropriate military specifications and standards are required by contract (e.g., MIL-I-6870 and MIL-STD-1530).

<u>Solution Shows Promise</u>?

Not Promising				Very Promising
1	2	3	4	5
		1	3	2

Weighted Mean= 4.23
Standard Score= 52.1

<u>Solution is Attainable</u>?

Not Feasible				Very Feasible
1	2	3	4	5
		1	3	2

Weighted Mean= 4.24
Standard Score= 57.9

<u>Estimated Cost of Solution</u>?

Very High				Very Low
1	2	3	4	5
		3	2	1

Weighted Mean= 3.69
Standard Score= 56.5

Promise + Feasibility = 8.47
Standard Score = 56.6

Promise + Feasibility + Cost = 12.16
Standard Score = 58.8

VI. DISCUSSION

The purpose of this research was to identify areas of concern and possible solutions with regard to Air Force NDI technician proficiency. Possible solutions were then consolidated into recommendations, which were evaluated for promise, feasibility, and cost. The recommendations presented in the preceding section should provide meaningful avenues to approach the identified concerns. Final decisions about which approaches should be attempted are, of course, left to the Air Force. The absolute ratings of promise and feasibility should provide decision makers with some guidance about priorities for the various strategies.

When evaluating the relative worth of a given recommendation, the reader is reminded that only the best identified strategies reached the panel for evaluation. In other words, the standard scores for promise, feasibility, etc., are measures of how that recommendation was rated relative to other highly selected recommendations. The weighted means provide a better estimate of how promising and feasible a recommendation was judged to be, and as the data indicate, most of the recommendations were judged to be fairly promising and feasible.

Finally, it should be noted that several of the recommendations can be combined to address more than one area of concern. The best example is the computerized trainer/evaluator, which contain features that would have an impact on a large number of concerns (e.g., initial training, OJT, performance measurement, feedback, standardization, and motivation). Although probably a relatively costly solution, such a solution could prove to be the most cost effective, because of the number of concerns it could address.

VII. REFERENCES

Ainsworth, L. (1985). Human factors considerations in NDI. 11th World Conference on Nondestructive Testing, 2, Dallas, Texas: Taylor Publishing Co., 1115-1119.

Alcorn, F. S., & O'Donnell, E. (1969). The training of nonphysician personnel for use in a mammography program. Cancer, 23, 879-884.

Annett, J. (1961). The role of knowledge of results in learning: A survey. (NAVTRADEVCEN Tech. Report No. 342-3), U. S. Naval Training Device Center, New York.

Annett, J. (1966). Training for perceptual skills. Ergonomics, 9, 459-468.

Berens, A. P., & Hovey, P. W. (December, 1981). Evaluation of NDE reliability characterization. AFWAL-TR-81-4160, Volume 1.

Bipes, T. U., Mullan, M. H., & Padronos, W. L. (1985). NDT technician training using the competency-based personalized instruction system. 11th World Conference on Nondestructive Testing, 3, Dallas: Taylor Publishing Co., 2078-2082.

Boisvert, B. W., Lewis, W. H., & Sproat, W. H. (1981). Uniform qualification of military and civilian nondestructive inspection personnel. Report No. LG81WP7254-003, Lockheed-Georgia Company, Marietta, Georgia.

Brock, J.E., Wells, R.G. & Abrams, M.L. (1974). Development and validation of an experimental radiograph reading training program. Report No. NPRDC-TR-74-33. Naval Personnel Research and Development Center, San Diego, California.

Buckner, D. N., & McGrath, J. J. (Eds.). (1963). Vigilance: A symposium. New York: McGraw Hill.

Burley, C. (1980). Comments Published in Mordfin, L.: Reliability of nondestructive evaluation. Critical Issues in Materials and Mechanical Engineering, 47, 133-147.

Bush, S. H., Becker, R. A., Cheng, C. Y., Collins, W. J., Crowley, B. R., Durr, J. P., Hazelton, W. S., Mathews, P. R., Muscara, J., Robinson, R. C., & Strosnider, J. (members of Pipe Crack Task Force). (1984). Report of the U.S. Nuclear Regulatory Commission piping review committee: Investigation and evaluation of stress corrosion cracking in piping of boiling water reactor plants. NUREG 1061, Volume 1, U.S. Nuclear Regulatory Commission, Washington, D.C.: U.S Government Printing Office.

Busse, L. J., et al. (1982). Characterization methods for ultrasonic test systems. NUREG/CR-2264, PNL-4215, U.S. Nuclear Regulatory Commission, Washington, D.C.: U.S Government Printing Office.

Chin Quan, H. R. (1974). Operator effects in NDI - part III, ultrasonics measurement on aluminum alloy blocks containing edge slots. Metallurgy Note 101, Australian Defense Scientific Service, Melborne: Aeronautical Research Laboratories.

Chin Quan, H. R., & Scott, J. G. (1977). Operator performance and reliability in NDT. In R. S. Sharpe (Ed.) Research techniques in nondestructive testing, Vol. III, London: Academic Press, 323-354.

Cowfer, C. D. (1984). Section XI ASME B&PV code, future trends, nondestructive examination requirements. The American Society of Mechanical Engineers, 84-PVP-42, 1-5.

Davies, D. R., & Parasuraman, R. (1982). The psychology of vigilance. London: Academic Press.

Davies, D. R., Shackleton, V. J., & Parasuraman, R. (1983). Monotony and boredom. In R. Hockey (Ed.), Stress and fatigue in human performance. Chicester, England: Wiley.

Davies, D. R., & Tune, G. S. (1970). Human vigilance performance. London: Staples.

Davis, J. M. (1988). Mathematical formulas and references for nondestructive testing: Ultrasonics. Itasca, Illinois: Technical Ink, Inc.

Davis, L. E., & Cherns, A. B. (1975). The quality of working life. London: Collier Macmillan.

Doctor, S. R., Deffenbaugh, J. D., Good, M. S., Green, E. R., Heasler, P. G., Mart, G. A., Simonen, F. A., Spanner, J. C., Taylor, T. T., & Van Fleet, L. G. (1987). Nondestructive examination (NDE) reliability for inservice inspection of light water reactors. NUREG/CR-4469, PNL-5711, Nuclear Regulatory Commission, Washington, D.C.: U.S Government Printing Office.

Drury, C. G., & Addison, J. L. (1973). An industrial study of effects of feedback and fault density on inspection performance. Ergonomics, 16, 159-169.

Eagle Technology Technical Report. (1985). Phase I, analysis of instructional content: Nondestructive inspection training. Eagle Technology, Inc., Orlando Florida.

Eitzen, D. G., Sushinsky, G. F., Chwirut, D.J., Bechtold, C.J., & Ruff, A. W. (April, 1975). Improved ultrasonic standard reference blocks. Report No. NBSIR 75-85, Washington, D.C.: National Bureau of Standards.

Embrey, D. E. (1979). Approaches to training for industrial inspection. Applied Ergonomics, 10(3), 139-144.

Enrietto, J. F. (1980). Comments published in L. Mordfin: Reliability of nondestructive evaluation. Critical Issues in Materials and Mechanical Engineering, 47, 133-147.

Fong, J. T. (1980). Comments Published in L. Mordfin: Reliability of nondestructive evaluation. Critical Issues in Materials and Mechanical Engineering, 47, 133-147.

Forsten, J., & Aaltio, M. (1982). On the detection of weld defects by x-ray and ultrasonic examination. British Journal of Non-Destructive Testing. 24(1), 33-46.

Gentner, D., & Stevens, A. L. (Eds.). (1983). Mental models. Hillsdale, New Jersey: Lawrence Erlbaum.

Glasch, K. J. (1987). Human reliability in nondestructive evaluation. Materials Evaluation, 45, 907-932.

Green, D. M., & Swets, J. A. (1974). Signal detection theory and psychophysics. Huntington, New York: Krieger.

Haines, N. F. (1977). The reliability of ultrasonic inspection. Proceedings of the international symposium on application of reliability technology to nuclear power plants.

Hanley, J. A., & McNeil, B. J. (1983). A method of comparing the areas under receiver operating characteristic curves derived from the same cases. Radiology, 148, 839-843.

Harris, D. H., & Chaney, F. B. (1969). Human factors in quality assurance. New York: Wiley.

Hellier, C. J. (1980). The myth of certification. Materials Evaluation, 38(9), 37-40.

Hellier C. J. (1984). Understanding certification - SNT-TC-1A, its benefits and problems, The American Society of Mechanical Engineers, 84-PVP-102, 1-3.

Herr, J. C. (1974). Human factors in NDE. Second ASM Materials Design Forum on Prevention of Structural Failure - The Role of Quantitative Nondestructive Evaluation, Port St. Luci, Florida.

Herr, J. C., & Marsh, G. L. (1978). NDT reliability and human factors. Materials Evaluation, 36(1), 41-46.

Holding, D. H. (1983). Fatigue. In R. Hockey (Ed.), Stress and fatigue in human performance. Chicester, England: Wiley.

Jamison, T. D., & McDearman, W. R. (1981). Studies of section XI ultrasonic repeatability. (Report NO. NP-1858), Palo Alto, California: Electric Power Research Institute.

Karimi, S. (1987). Human factors affecting NDE technician performance. Report Prepared for Electric Power Research Institute, RP1570-19, Palo Alto, CA.

Kemmler, E. (1980). Comments Published in L. Mordfin: Reliability of nondestructive evaluation. Critical Issues in Materials and Mechanical Engineering, 47, 133-147.

Kino, G. S. (1980). Comments Published in L. Mordfin: Reliability of nondestructive evaluation. Critical Issues in Materials and Mechanical Engineering, 47, 133-147.

Lautzenheiser, C. E., Trigilio, R. F., & Meredith, W. R. (1979). NDE personnel needs in nuclear industry. American Nuclear Society Transactions, 32, 98.

Lewis, W.H., Sproat, W.H., Dodd, B.D., and Hamilton, J.M. (1978). Reliability of nondestructive inspections: Final report. Report No. SA-ALC/MME 76-6-38-1, Marietta, Georgia: Lockheed-Georgia Company.

Lusted, L. V., Bell, R. S., Edwards, W., Roberts, H. C., & Wallace, D. L. (1977). Evaluating the efficacy of radiologic procedures by Bayesian methods. In K. Snapper (Ed.), Models in metrics for decision makers. Washington, D.C.: Information Resources Press.

McMaster, R. C. (1984). The costs of immortality in nondestructive testing. Materials Evaluation. 42, 402-403.

Michalski, F., Stadthaus, M., & Kaiser, D. (1976). The human eye, an instrument for nondestructive inspection. Proceedings of the Eighth World Conference on Nondestructive Testing, Cannes, France, Section 3A3, 1-7.

Mitchell, D. K., & Meister, R. P. (1973). A technique for selectively enhancing the ultrasonic flaw response over accompanying noise. Proceedings of the 9th Symposium on Non Destructive Evaluation, San Antonio, 29-37.

Mordfin, L. (1980). Reliability of nondestructive evaluation. Critical Issues in Materials and Mechanical Engineering, 47, 133-147.

Nakatsuji, T., Kuramochi, M. & Fujimori, T. (1983). Defect measuring accuracy and defect detection probability by manual ultrasonic examination. Shimizu Technical Research Bulletin. 2, 37-44.

Packman, P. F., & Malpani, J. K. (1978). Reliability of critical components after NDI inspection. Paper presented at the ASNT Spring Conference.

Peddada, T., & Bennett. C. A. (1984). Inspection contrasting self-pacing and machine pacing. Proceedings of the human factors society 28th annual meeting, The Human Factors Society: Santa Monica, California, 675-677.

Petru, J. A. (1985). U.S. Air Force reliability programs. 11th World Conference on Nondestructive Testing, 2, Dallas, Texas: Taylor Publishing Co., 1120-1123.

Posakony, G. J. (1985). Influence of the pulser on the ultrasonic spectrum: The results of an experiment. Materials Evaluation, 43(4), 413-419.

Posner, M. I., & Keele, S. W. (1968). On the genesis of abstract ideas. Journal of Experimental Psychology, 77, 353-363.

Rain, C. (1984). Uncovering hidden flaws. High Technology, 49-55.

Ramsey, J. D. (1983). Heat and cold. In A. Hockey (Ed.), Stress and fatigue in human performance. Chicenter, England: Wiley.

Reinhardt, E. R. (1977). A study of in-service ultrasonic inspection for BWR piping welds. Report no. NP-436SR, Palo Alto, California: Electric Power Research Institute.

Rummel, W. D. (1983). Human factors considerations in the assessment of nondestructive evaluation (NDE) reliability. In D. O. Thompson & D. E. Chimenti (Eds.), Review of progress in quantitative nondestructive evaluation, New York: Plenum Press, 37-46.

Spanner, J. C., Badalamente, R. V., Rankin, W. L., & Triggs, T. J. (1986). Human reliability impact on inservice inspection: Phase 1 summary report. NUREG/CR-4436, PNL-5641, BHARC-400/85/016, Nuclear Regulatory Commission, Washington, D.C.: U.S Government Printing Office.

Sproat, W.H. (Undated Report Provided by AF NDI Program Office). Elements of NDE reliability data acquisition as related to human factors. Lockheed-Georgia Report.

Sproat. W. H., & Sharp, H. (1979). Measurement of NDI technician proficiency on C-5 wing hot spot inspections. Lockheed Report LG 79 ER0161.

Stone, D. E. W. (1983). Comment Recorded by W. D. Rummel and reported in: Human factors considerations in the assessment of nondestructive evaluation (NDE) reliability. In D. O. Thompson & D. E. Chimenti (Eds.), Review of progress in quantitative nondestructive evaluation, New York: Plenum Press, 37-46.

Stroh, C. M. (1971). Vigilance: The problem of sustained attention. New York: Pergamon.

Summers, R.H. (1984). Nondestructive inspection: Improved capabilities of technicians. Air Force Human Resources Laboratory Final Report, AFHRL-TP-83-63, Brooks Air Force Base, Texas: Air Force Systems Command.

Swets, J. A. (1977). Signal detection theory applied to vigilance. In R. R. Mackie (Ed.), Vigilance: Relationship among theory, physiological correlates and operational performance. New York: Plenum.

Swets, J. A. (1983a). Assessment of NDT systems - part I: The relationship of true and false detections. Materials Evaluation, 41(11), 1294-1298.

Swets, J. A. (1983b). Assessment of NDT systems - part II: Indices of performance. Materials Evaluation, 41(11), 1299-1303.

Swets, J. A., & Pickett, R. M. (1982). Evaluation of diagnostic systems: Methods from signal detection theory. New York: Academic Press, Inc.

Swets, J. A., Pickett, R. M., Whitehead, S. F., Getty, D. J., Schnur, J. A., Swets, J. G., & Freeman, B. A. (1979). Assessment of diagnostic technologies. Science, 205, 753-759.

Taylor, T. T., & Selby, G. P. (1981). Evaluation of ASME section XI reference level sensitivity for initiation of ultrasonic inspection examination. Report No. NUREG/CR-1957, U.S. Nuclear Regulatory Commission, Washington, D.C.: U.S Government Printing Office.

Triggs, T.J., Rankin, W.L., Badalamente, R.V., & Spanner, J.C. (1986). Human reliability impact on inservice inspection: Review and analysis of human performance in nondestructive testing (emphasizing ultrasonics). NUREG/CR-4436, PNL-5641, BHARC-400/85/016, Vol.2, Nuclear Regulatory Commission, Washington, D.C.: U.S Government Printing Office.

Ward, J. M. (1985). A total quality improvement programme. In P. A. McKeown, (Ed), Proceedings of the 7th international conference on automated inspection and product control, IFS (Publications) Ltd. & North-Holland, 1-12.

Watkins, B., & Cowburn, K. J. (1980). The reliability of defect detection by ultrasonics with particular reference to the ISI of nuclear pressure vessels. United Kingdom Atomic Energy Authority Report No. ND-R-420(4), Warrington, England: Risley.

Watson, C. S., & Nichols, T. L. (1976). Detectability of auditory signals presented without defined observation intervals. Journal of the Acoustical Society of America, 59, 655-668.

Weiner, E. L. (1968). Training for vigilance: Repeated sessions with knowledge of results. Ergonomics, 11, 547-556.

Wheeler, W. A., Rankin, W. L., Spanner, J. C., Badalamente, R. V., & Taylor, T. T. (1986). Human factors study conducted in conjunction with a mini-round robin assessment of ultrasonic technician performance. NUREG/CR-4600, PNL-5757, BHARC-400/86/001, RX, R5, Nuclear Regulatory Commission, Washington, DC: : U.S Government Printing Office.

Whittle, M. J., & Coffey, J. A. (1981). The PISC exercise: A discussion of its relevance to ultrasonic inspection of pressure vessels. British Journal of Nondestructive testing, 23(2), 71.

Wickens, C. D. (1984). Engineering psychology and human performance. Columbus, Ohio: Merrill.

Williges, R. C., & North, R. A. (1972). Knowledge of results and decision making performance in visual monitoring. Organizational Behavior and Human Performance, 8, 44-57.

Winslow, R. (1984, November 7). Quality quandary: Regulators investigate harassing of inspectors at new nuclear plants. Wall Street Journal (Midwest edition), Vol. LSV (NO. 18).

Yonemura, G. T. (1981). Visual acuity testing of radiographic inspectors in nondestructive inspection. NBS Technical Note 1143, Washington D.C.: National Bureau of Standards.

Zong, R. (1980). Comments Published in L. Mordfin: Reliability of nondestructive evaluation. Critical Issues in Materials and Mechanical Engineering, 47, 133-147.